2020

Smad4 With R361 Hotspot Mutations Retains The Ability To Bind To Lef1 And Boosts Wnt Signaling In Colorectal Cancer Cells

Claudia B. Lanauze
University of Pennsylvania

Smad4 With R361 Hotspot Mutations Retains The Ability To Bind To Lef1 And Boosts Wnt Signaling In Colorectal Cancer Cells

Abstract

Colorectal cancer (CRC) is one of the leading causes of deaths in the United States. Roughly 150,000 new cases are diagnosed each year, resulting in ~50,000 deaths. About 10-30% of CRC patients harbor either loss of or missense mutations in SMAD4, a critical component of the TGF-β signaling pathway. Our lab and others have shown that complete loss of Smad4 results in the increase in tumor size, microvascular density, and frequency of metastasis in CRC xenograft models. While the role of Smad4 loss in CRC progression has been extensively studied, the pathophysiological function of missense mutations in Smad4 is not fully understood. At the molecular level, these mutations usually map to the MH2 domain and eliminate residues that are involved in the formation of the heteromeric complex with regulatory Smads (R-Smads) such as Smad2/3 and ensuing transcriptional activation. These detrimental effects suggest that SMAD4 missense mutations can be categorized as loss-of-function (LOF). However, uncharacteristically for LOF mutations, they cluster in a few hotspots (e.g., R361), which is more consistent with a gain-of- or neomorphic function. Here, we investigated the functional role of Smad4 R361 mutants in vitro by re-expressing two R361 Smad4 variants in Smad4-null CRC cells. As predicted, R361 mutations disrupted Smad2/3-Smad4 heteromeric complex formation and abolished canonical TGF-β downstream signaling. In that, they were similar to SMAD4 loss. However, RNA sequencing and subsequent RT-PCR revealed that Smad4mut cells possess the known gene signature associated with enhanced LEF1 protein function and increased WNT signaling. Mechanistically, Smad4 mutant proteins retained binding to LEF1 protein and directed a commensurate increase in downstream Wnt signaling as measured by TOP/FOP luciferase assay. Consistent with these findings, human CRCs with SMAD4 missense mutations were less likely to acquire activating mutations in the key Wnt pathway gene CTNNB1 (encoding beta catenin) than CRCs with truncating SMAD4 nonsense mutations. The former was also associated with shorter survival. Collectively, these studies implicate a TGF-β ligand-independent gain of function role for mutant Smad4 in CRC.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Cell & Molecular Biology

First Advisor

Andrei Thomas-Tikhonenko

Keywords

Cancer, Colorectal Cancer, SMAD4, TGFbeta

Subject Categories

Cell Biology | Medicine and Health Sciences

SMAD4 WITH R361 HOTSPOT MUTATIONS RETAINS THE ABILITY TO BIND TO LEF1 AND BOOSTS WNT SIGNALING IN COLORECTAL CANCER CELLS

Claudia B. Lanauze Torres

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

_____

Andrei Thomas-Tikhonenko

Mildred L. Roeckle Endowed Chair in Pathology; Chief, Division of Cancer Pathobiology
Professor of Pathology & Laboratory Medicine and Pediatrics

Graduate Group Chairperson

_____

Daniel S. Kessler

Associate Professor of Cell and Developmental Biology

Dissertation Committee

Christopher J Lengner, PhD (chair), Associate Professor, Department of Biomedical Sciences
Brian Keith, PhD, Associate Investigator, Abramson Family Cancer Research Institute
Ellen Puré, PhD, Professor of Pharmacology
Anil Rutgsi, MD, Director of the Herbert Irving Comprehensive Cancer Center, Associate Dean of Oncology, Chief of Cancer Services

This thesis is dedicated to my parents and my siblings, for their endless love and support.

ACKNOWLEDGMENTS

I think it is safe to say that I would not have been able to start, continue, and finish my journey without help from a lot of people in some way or the other. I am grateful for my advisor, Andrei Thomas-Tikhonenko, and his mentorship over the years. Thank you for having faith in me when I joined your lab, for guiding me and giving me the freedom to explore new ideas and interests.

I would like to thank the current and former members of the ATT lab who have helped and supported me throughout the years. I would especially like to thank Priyanka Sehgal and Colleen Harrington. Not only have they been amazing coworkers, but I am so lucky to also call them my friends. Priyanka, thank you for helping me first-hand with experiments and thinking about my project when it didn't make sense, I would not have been able to finish this degree without all of your help and input. Colleen, thank you for being the ray of sunshine and an incredible friend to me over the past years. I can honestly say I could not have gotten through graduate school without you guys.

Thank you to the CAMB program, especially Dan Kessler, Kathy O'Connor-Cooley, Christina Strathearn, Meagan Schofer, and Anna Kline for their constant efforts to provide the best environment for Penn students to thrive. To Arnaldo Diaz, for allowing me to come to Penn and taking a chance on me when I did not think anyone else would. Lastly, to the CTRB 4th floor and Cancer Pathobiology Department, thank you for your support, feedback, and post-6PM conversations.

I am eternally grateful for the close friendships I have made throughout my time at Penn. Rizwan Saffie, Dahmane Ouazia, Izmarie Poventud-Fuentes and Glendon Wu. Thank you for your support, friendship and providing a shoulder to cry on. I can't wait to continue our happy hours when it is safe to travel again. To my family, thank you for all your support throughout the years. Even though I am not physically close to them, they

somehow manage to lift my spirits and make me laugh all the time. “Just take a plane and fly down to Puerto Rico, that will make you feel better.” To my parents, José Lanauze and Edith Torres, who have sacrificed so many things for me to pursue my goals. To my siblings, my uncle, my aunt, and my cousin: Gabriela, Javier, Tio Tongui, Tigi Gigi, and Mónica, for being my cheerleaders throughout this entire journey. And to my grandparents, who continuously put me in their thoughts and prayers. I feel blessed to have such a supportive family.

And finally, to my husband Bradley Browning, for continuously pushing me to be the best version of myself. I am eternal grateful for his patience and support these past couple of years. I am so lucky to have him by my side, and I can’t wait to see what the future has in store for us.

ABSTRACT

SMAD4 WITH R361 HOTSPOT MUTATIONS RETAINS THE ABILITY TO BIND TO LEF1 AND BOOSTS WNT SIGNALING IN COLORECTAL CANCER CELLS

Claudia B Lanauze Torres

Andrei Thomas-Tikhonenko

Colorectal cancer (CRC) is one of the leading causes of deaths in the United States. Roughly 150,000 new cases are diagnosed each year, resulting in ~50,000 deaths. About 10-30% of CRC patients harbor either loss of or missense mutations in SMAD4, a critical component of the TGF-β signaling pathway. Our lab and others have shown that complete loss of Smad4 results in the increase in tumor size, microvascular density, and frequency of metastasis in CRC xenograft models. While the role of Smad4 loss in CRC progression has been extensively studied, the pathophysiological function of missense mutations in Smad4 is not fully understood. At the molecular level, these mutations usually map to the MH2 domain and eliminate residues that are involved in the formation of the heteromeric complex with regulatory Smads (R-Smads) such as Smad2/3 and ensuing transcriptional activation. These detrimental effects suggest that SMAD4 missense mutations can be categorized as loss-of-function (LOF). However, uncharacteristically for LOF mutations, they cluster in a few hotspots (e.g., R361), which is more consistent with a gain-of- or neomorphic function. Here, we investigated the functional role of Smad4 R361 mutants *in vitro* by re-expressing two R361 Smad4 variants in Smad4-null CRC cells. As predicted, R361 mutations disrupted Smad2/3-Smad4 heteromeric complex formation and abolished canonical TGF-β downstream signaling. In that, they were similar to SMAD4 loss. However, RNA sequencing and subsequent RT-PCR revealed that Smad4mut cells possess the known gene signature associated with enhanced LEF1 protein function and increased WNT signaling. Mechanistically, Smad4

mutant proteins retained binding to LEF1 protein and directed a commensurate increase in downstream Wnt signaling as measured by TOP/FOP luciferase assay. Consistent with these findings, human CRCs with SMAD4 missense mutations were less likely to acquire activating mutations in the key Wnt pathway gene CTNNB1 (encoding β-catenin) than CRCs with truncating SMAD4 nonsense mutations. The former was also associated with shorter survival. Collectively, these studies implicate a TGF-β ligand-independent gain of function role for mutant Smad4 in CRC.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: GENERAL INTRODUCTION

OVERVIEW

Colorectal Cancer (CRC) is the second most common cause of cancer death in the United States. Early-stage CRC is successfully treated by surgical intervention; however, relapse can occur even if the primary tumor was successfully removed during the initial treatment. Reemergence of colorectal cancer usually leads to metastasis, often resistant to the current chemotherapy regimen. Patients diagnosed with localized disease have about a 90% 5-year survival rate, compared to those diagnosed with metastatic CRC, which have only 14% 5-year survival rate[1]. Therefore, elucidating molecular events that lead to the advancement of CRC could open the door for future targeted therapies for such patients.

About 10-30% of CRC patients harbor some loss of Smad4, a critical component of the TGF-β signaling pathway, either by deep deletion or missense mutations[2]. Based on limited research, naturally occurring Smad4 mutations are thought to be loss-of-function. However, there is a big gap in our understanding of the role of mutant Smad4 in TGF-β signaling in CRC progression or any other acquired novel function, if any. Here, we provide relevant background information important to understand potential role of TGF-β signaling to CRC progression.

COLORECTAL CANCER (CRC)

In 2020, almost 148,000 individuals will be diagnosed with CRC and over 53,000 will die from the disease. The majority of CRC cases occur in individuals 50 years of age or older, however, up to 12% of newly diagnosed CRC cases are attributed to individuals younger than 50 years[1]. Although 10% of patients with CRC have a genetic predisposition to the disease, the majority of CRC cases are sporadic.

Benign GI tumors are a varied group but localized lesions that project above the surrounding tissue are commonly termed polyps. Although some polyps (termed hyperplastic polyps – polyps less than 5mm in size) are not a major precursor of CRC, the adenomatous polyp (adenoma) is thought to be the most important precursor lesion[2]. Even though only a fraction of adenomas will progress to cancer, there is a risk of CRC in individuals whose adenomas are not removed, and polypectomy decreases the risk of CRC [3,4].

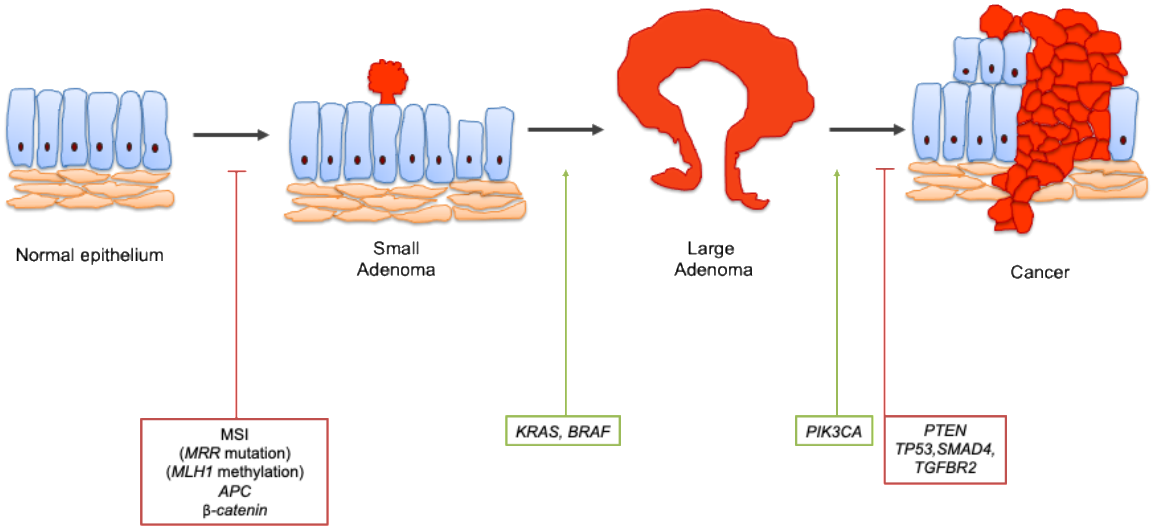

Figure 1. Molecular Events in CRC

In the 'classic' CRC formation model originally developed by Dr. Bert Vogelstein (Figure 1), CRC pathogenesis occurs due to the progressive accumulation of genetic and epigenetic alterations. The loss of genomic stability accelerates the accumulation of mutations in oncogenes and tumor suppressor genes. This creates a clonal growth advantage that leads to the outgrowth of malignant cells, which can ultimately develop into invasive adenocarcinoma[5–7]. Although some specific genetic events have been described as initiating events for CRC, the progressive accumulation of the alterations, rather than the order, seems to be key for the progression of colorectal tumors[6].

At least two major genomic pathways are involved in the origin and progression of CRC: microsatellite instability (MSI) and chromosomal instability (CIN). About 15% of CRCs occur due to microsatellite instability (MSI), a genetic event that causes hyper mutability in the DNA as a result in loss of functional DNA mismatch repair genes (MMR)[8]. Most commonly, the promoter of *hMLH1* will be hypermethylated, resulting in gene silencing[9]. In some instances, MSI can occur at protein-coding sequences that eventually result in point or frame-shift mutations that will yield a truncated, nonfunctional protein[10]. In these tumors, frequent inactivating mutations (~85% of colorectal tumors) occur to the transforming growth factor beta receptor 2 (TGFβRII)[11] gene, a critical component of the transforming growth factor beta (TGF-β) signaling family.

Chromosomal instability (CIN) represents 85% of sporadic CRCs[12] and is characterized by allelic losses, amplifications, and translocations; usually associated with mutations in tumor suppressor genes or oncogenes[13]. This process involves the alteration of several pathways including the WNT pathway through the loss of APC, the addition of constitutive active mutations in KRAS or BRAF, mutations in the PI3K pathway, and inactivation of SMAD4 and TP53. Specifically:

1. **Adenomatous polyposis coli (*APC*) gene and WNT signaling pathway:** Approximately 80% of all sporadic colorectal adenomas and carcinomas have somatic mutations that inactivate the tumor suppressor APC, a major binding partner and regulator of β-catenin in the canonical WNT signaling pathway[3,14]. Briefly, APC binds and collaborates with other proteins to form the β-catenin destruction complex, which ubiquitinates and targets β-catenin for proteasomal degradation[15]. *APC* inactivation will primarily result in the stabilization of β-catenin, which mimics constitutively active WNT signaling[16].
2. **Mutations affecting the TGF-β pathway components in CRC:** loss of heterozygosity (LOH) of chromosome 18q is observed more than 70% of CRCs[17].

Two chromosome 18q tumor suppressor genes, *SMAD2* and *SMAD4,* encode proteins that function downstream of the TGF-β receptor complex[18,19]. Mutations that inactivate *SMAD4* are found in ~10–15% of CRCs, and SMAD2 inactivating mutations are found in ~5% of CRCs. Mutations inactivating the *SMAD3* gene, which maps to chromosome 15, are also only found in ~5% of CRCs[3,20–22].

3. **Alterations in the KRAS, BRAF, PIK3CA and PTEN signaling pathway:** Ras family is comprised of a group of small GTP-binding proteins that regulate different signaling pathways involved in cell proliferation, differentiation, migration, survival and apoptosis[23]. Active Ras can bind to a variety of effector proteins, the best characterized being Ras kinases[24] and phosphatidylinositol 3-kinases[25] (PI3-K). It's three members, KRAS, HRAS and NRAS are common targets for somatic mutations in human cancers[23]. *KRAS* somatic mutations are found in up to a half of CRC patients. The vast majority of mutations affect codon 12 and contribute to colorectal adenoma development[17,26]. Mutations of *RAS* are clinically relevant given that their presence interferes with the response of monoclonal antibodies against epidermal growth factor receptor (EGFR) like cetuximab and panitumumab[27,28].

Six research groups identified a consensus for gene expression based on a subtyping classification system for CRC, which resulted into four consensus molecular subtypes (CMS). The four CMS subtypes identified (CMS1- CMS4) differ in genetic and epigenetics, as well as the signaling pathways they follow. While CMS1 is characteristic of tumors with MSI, tumors with CIN can be subclassified based on gene expression: CMS2 (canonical subtype), CMS3 (metabolic subtype) and CMS4 (mesenchymal subtype)[29].

Around 14% of all CRCs are classified into the CMS1 subtype, from which the majority of patients (88%) have inherited disease and remaining 12% are sporadic (non-inherited) CRCs[29]. They are also defined by having high *BRAF V600E* mutation rate[30] and a widespread hypermethylation status, resulting in loss of tumor suppressor function. Out of the four subgroups, CMS1 tumors show marked upregulation of proteins involved in immune response pathways[12,29].

The CMS2 (canonical) subtype accounts for the largest number of CRC patients, with about 39%[29]. There are more frequent copy number losses in tumor suppressor genes and copy number gains in oncogenes in CMS2 compared to other subtypes. Thought to arise from the canonical adenoma-to-carcinoma pathway, the gene expression profile in CMS2 is often characterized by having hyperactive Myc and WNT/β-catenin signal transduction pathways[29,31]. CMS3 subtype, also known as the metabolic subtype, accounts for 37% of CRCs. Although KRAS mutations are present in every CRC subtype, they are significantly enriched in this subtype. CMS3 displays prominent metabolic activation, a hallmark that can prove useful for specific future targeted therapies[31]. Finally, CMS4 is defined by having microsatellite stable (MSS) CRC tumors. Tumors show upregulation of genes involved in EMT and it has signatures associated with the activation of TGF-β signaling. In addition, samples from CMS4 subtype have significant overexpression of proteins involved in angiogenesis, matrix remodeling pathways and the complement-mediated inflammatory system[29,32]. Overall, these specific subtypes can help determine treatment strategies for individual CRC patients, based on the mutation panel and activated pathways in their tumors.

TGF-β and WNT pathways do not work in isolation, as there are examples of known interactions between them in cancer. β-catenin, the main effector of WNT signaling, has been found to form complexes with Smad proteins to enhance β-catenin protein stability and facilitate its transcriptional activity in chondrocytes[33]. The WNT effector LEF1

has been shown to co-occupy the CDH1 (E-cadherin) promoter with SMAD proteins, providing a molecular mechanism for the observed cooperation of WNT and TGF-β in EMT[34]. For the next sections of this introduction, we focus on describing TGF-β and WNT signaling pathways and their deregulation in colorectal cancer, as it pertains to the understanding of this study.

TRANSFORMING GROWTH FACTOR β (TGF-β) SUPERFAMILY SIGNALING OVERVIEW

The transforming growth factor beta (TGF-β) is a family of secreted polypeptide growth factors which include TGF-βs, activins, nodals, bone morphogenic proteins (BMPs) and others that play an important role in a variety of cellular processes such as cell growth, differentiation, migration and apoptosis[35].

The TGF-β signaling pathway is activated by TGF-β ligands (TGFβ1, β2, and β3), which bring together receptor type I and type II (RI and RII), both serine/threonine kinase receptors. Upon receptor dimerization, type II receptors phosphorylate and activate type I receptors. Following receptor phosphorylation and activation, type I receptor kinase will directly phosphorylate the receptor-regulated Smads (R-Smads) on the last two serine of a conserved specific phosphorylation motif, SSXS, located at the extreme carboxyl terminus of the MH2 domain[36].

The TGF-β ligand will phosphorylate Smad2 and 3 while BMP ligand and receptors will activate Smad1, 5 and 8. Once activated, R-Smads will form a complex with Smad4, also known as common mediator Smad. Once the complex forms, it will translocate into the nucleus, bind to DNA and other transcription factors where it will modify a different set of target genes[37,38].

TGFβRII receptors not only induce Smad-mediated responses but also activate Smad-independent responses, which allow additional versatility and diversification of TGF-β family responses. The non-Smad pathways include, but are not limited to, the activation of other mediators such as those involved in the Erk, JNK, and p38 MAPK kinase pathways[39] (Figure 2).

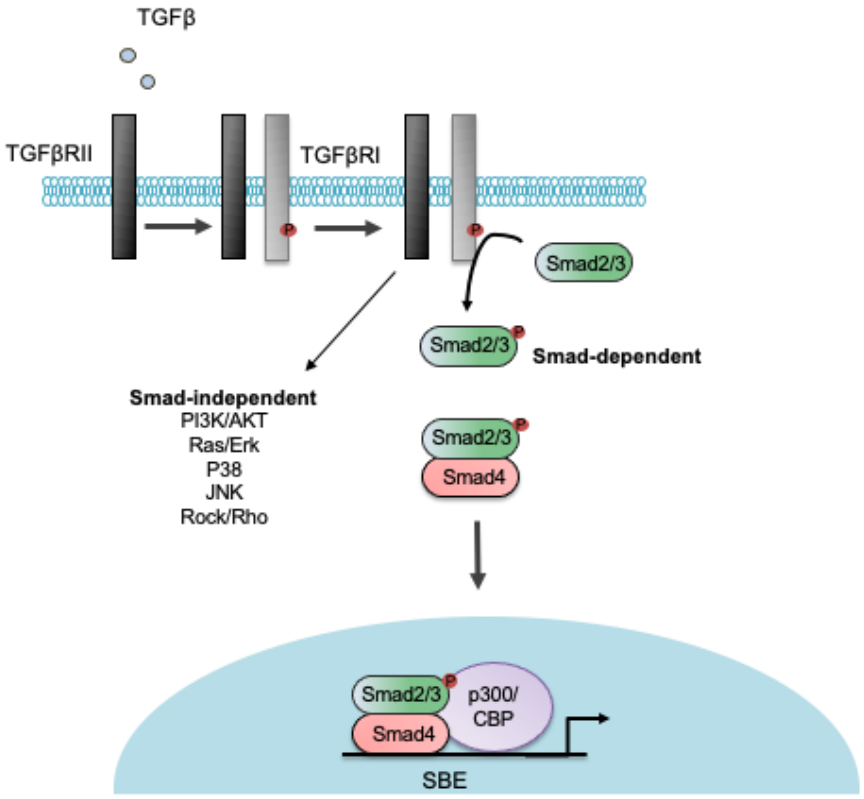

Figure 2. Smad-dependent and Smad-independent TGF-β signaling

SMAD PROTEINS AS EFFECTORS OF TGF-β

The eight-known Smads are around 500 amino acids in length and are divided into three classes according to their function. The receptor regulated Smads, or R-Smads, are direct substrates corresponding to their receptors. As mentioned above: Smad1, 5, and 8 serve principally as substrates for BMP receptors, and Smad2 and 3 for the TGF-β

receptors. The common mediator Smad (co-Smad), Smad4, participates in TGF-β by binding to R-Smads. Lastly, there are inhibitor Smads or I-Smads, which includes Smads 6 and Smad7: they antagonize TGF-β signaling by interfering with Smad-receptor or Smad-Smad interactions[35]. Smad7 can inhibit both TGF-β and BMP, whereas Smad6 preferentially inhibits BMP signaling[40,41].

Smad proteins are composed of two globular domains, joined together by a proline-rich linker region. The Mad homology 1 (MH1) domain, located at the N-terminus, is highly conserved among all R-Smads and Smad4, but not Smad6 and 7. The MH1 domain, which contains a β-hairpin structure, functions as a DNA-binding domain. The most abundant isoform of Smad2 contains a 30-amino acid insertion encoded in exon 3, which is thought to displace the β-hairpin loops, blocking its ability to bind to DNA[42,43]. A splice variant of Smad2 which lacks exon 3 (Smad2Δ3) binds to DNA equivalently to Smad3[44]. Smad3 and Smad4 recognize a specific sequence in the DNA called the Smad-binding element (SBE), which is defined as GCTC or its reverse complement, AGAC[43].

The MH1 domain is followed by a linker region, which contains several important peptide motifs including phosphorylation sites for several classes of protein kinases including MAPKs and CDKs[35,45]. R-Smads and I-Smads also contain a conserved proline-tyrosine (PY) motif that mediates interaction with the WW domains in the Smad-interacting proteins Smurf1 and Smurf 2. Smurfs are E3 ubiquitin ligases of the C2-WW-HECT domain class that catalyze ubiquitin-mediated degradation of certain Smads and Smad-associated proteins, including nuclear oncoprotein SnoN and the TGF-β-receptor complex[46–48]. The linker region in Smad4, however, contains a CRM1 recognized nuclear export signal (NES)[49] and also part of the Smad activation domain (SAD)[50], which is required for transcriptional activation.

The Smad MH2 domain is highly conserved among all Smads and is responsible for receptor recognition in the case of R-Smads, transactivation, interaction with

transcription factors and Smad-Smad oligomerization[51–53] (Figure 3). A crystal structure for Smad4 revealed that Smad4 hetero-oligomerizes with Smad2 through a defined L3 loop in the MH2 domain[53] and several Smad-inactivating mutations found in tumors map to the MH2 domain interface and inhibit Smad oligomer formation[54], including mutation hotspot region in CRC in Smad4 (Asp351-Pro356, Arg36) that is involved in binding to R-Smads[55].

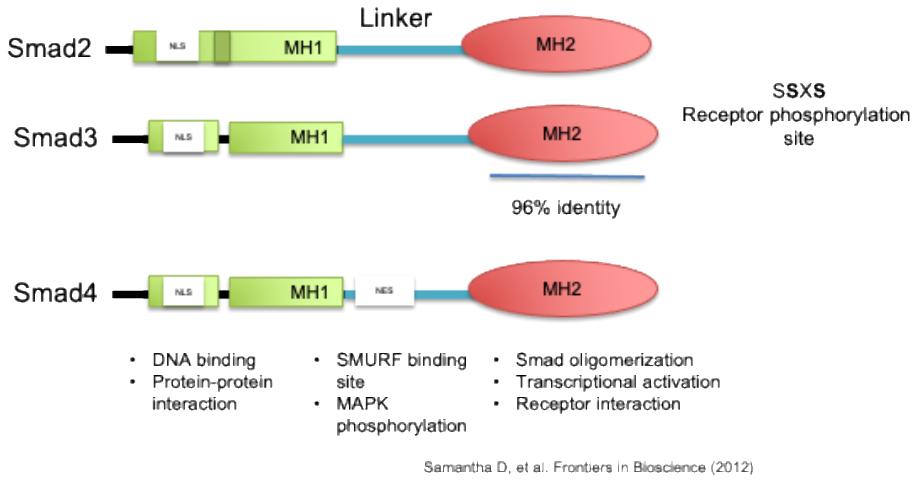

Figure 3. Smad proteins and their structural elements

As stated above, activated Smad complexes accumulate in the nucleus, where they bind DNA directly or indirectly via other transcription factors and then regulate expression of different target genes. Smads interact with co-activators including p300 and CBP which have histone acetyl transferase (HAT) activity[56]. This interaction is mediated by the Smad activation domain in R-Smads and Smad4 (located within the linker and MH2 domain) and is essential for maximal transcriptional activation[57]. Smads can also bind proteins that will repress Smad-dependent transcription, most famously Ski and SnoN proteins. Because the Ski binding surface on Smad4 overlaps with that required for binding of R-Smads, Ski and R-Smads compete for mutual exclusive binding to Smad4, leading to disruption of functional Smad4-R-Smad complexes[58]. Overall, binding of Smads to specific transcription factors allow for a variation in TGF-β functions.

COMMON MEDIATOR SMAD – SMAD4

As previously mentioned, Smad4 is referred to as the common mediator Smad due to it being a central mediator for both TGF-β and BMP signaling[59]. It is composed of 552 amino acids (~60kDA) and its primary structure consists of three major parts, including the N-terminal MH1 domain, the C-terminal MH2 domain and the linker region in between them[35]. Smad4 forms a complex with Smad2 and Smad3, and translocate to the nucleus where it activates transcription of multiple TGF-β response genes[35,36,49,53]. Smad4 is an important component of the TGF-β signaling pathway and plays a critical role in its tumor suppressive role, mainly by inducing cell cycle arrest and apoptosis. For example, Smad4 can induce p21CIP expression by binding to its promoter, resulting in inhibition of cell proliferation[60,61].

TGF-β SIGNAING IN CANCER

TGF-β signaling plays a dual role in cancer progression. In normal and premalignant cells, TGF-β acts as a tumor suppressor by inhibiting cell growth and inducing apoptosis. For example, Smad transcriptional complexes can target p15INK4b and p21CIP[62,63] for transcriptional activation leading to cyclin-dependent kinase (CDK) inhibition and subsequent cell cycle arrest. Interestingly, the proto-oncogene c-Myc can block TGF-β induced growth arrest[64]. Smad3 can directly bind to a TGF-β inhibitory element (TIE) in the promoter of *c-myc* and repress its transcription[65], leading to subsequent downregulation of c-Myc and upregulation of p21CIP[66] and p15INK4b[67].

Even though TGF-β/Smad signaling can inhibit cell proliferation and induce apoptosis in non-cancerous cells, TGFβ1 mRNA is highly expressed within tumor tissues of many cancer types[68]. For example, CRC patients with high levels of TGF-β are associated with a high risk of cancer recurrence compared to those with low levels of TGF-

β in their tumors[69]. This is because during the later stages of cancer, TGF-β acts as a tumor promoter by inducing tumor invasion and metastasis, which can be done through modulation of the surrounding tumor microenvironment. As an illustration, when CRC-derived cell lines are engineered to secret active TGFβ1, they display increased metastasis in mice to both liver and lung. This event was found to be mainly driven by TGF-β activation in tumor-associated stromal cells[69].

An important effect of TGF-β on tumor progression is the induction of epithelial to mesenchymal transition (EMT), wherein epithelial-like tumor cells generate mesenchymal-like characteristics, including decreased cell-cell adhesion and increased migration, invasion, and apoptosis-resistant properties[70,71]. TGF-β, among other tumor-associate growth factors, is responsible for the induction of multiple EMT-inducing factors including Snail, Slug, ZEB and Twist[70]. TGF-β regulates EMT through Smad-dependent and Smad-independent pathways, such as PI3K/Akt, ERK1/2, p38 and MAPK. For example:

1. **Smad-dependent pathway EMT:** Downregulation of Smad2 or SARA (Smad anchor for receptor activation; an endocytic adaptor protein that facilitates Smad2 binding to the TGFβRI and phosphorylation) increases the induction of EMT in human renal epithelial cells[72]. Also, ubiquitination of Smad4 by the E3 ubiquitin ligase TIF1γ disrupted nuclear Smad complex, antagonizing TGF-β-induced EMT in human mammary epithelial cells[73].
2. **Smad-independent pathway EMT:** On the other hand, TGF-β receptors can directly influence EMT via post-translational modifications of regulators of epithelial cell polarity and tight-junction assembly as is the case with Par6 protein. TGFβRII will phosphorylate Par6, which in turn recruits Smurf1, an E3 ubiquitin ligase. Smurf1 will then target RhoA for ubiquitination and degradation, resulting in loss of tight junctions and cell polarity[74,75]. Recent studies have also shown that TGF-β activates mTOR pathway during EMT via

PI3K and Akt[76]. The activation of mTOR, in turn, contributes to increased mRNA expression of transcription factor Snail which leads to changes in migration and invasion[77].

Among other processes that TGF-β can control is angiogenesis, the process by which new blood vessels form from pre-existing blood vessels. Tumors can co-opt this mechanism and use it to their advantage, allowing for increased nutrients and oxygen supply[78]. For many tumors, increased vascularity can result in metastasis as the new blood vessels provide a route by which tumors cells exit the primary tumor site and enter the circulation[79]. Vascular density can provide a prognostic indicator of metastatic potential, with the highly vascular primary tumors having higher incidence of metastasis than poorly vascularized tumors[80]. For example, our lab has shown that downregulation of Smad4 protein in colorectal cancer cells results in increased angiogenesis *in vivo,* due to downregulation of antiangiogenic factor CTGF[81].

ALTERATIONS OF SMAD4 IN CANCER

Malignant cells can circumvent the suppressive effects of TGF-β either through inactivation of core components of the pathway, such as the receptors or by downstream alterations that disable the tumor suppressive arm the pathway[37]. Smad4 is an essential signal transducer of the transforming growth factor β (TGF-β) signaling pathway and is described as a tumor suppressor. Loss of Smad4 expression promotes malignancy in colorectal cancer tumors cells[82] and serves as a prognostic marker in colorectal cancer patients[83–85].

Furthermore, Smad4 mutations are associated with the occurrence of JPS[86] and mutations or deletions have been found in about half of pancreatic cancers as well as

approximately 10-30% of colorectal cancers[2,87]. In these tumors, Smad4 mutations appear more frequently in the MH2 domain and may be missense or nonsense mutations[88]. The reported nonsense mutation found in PDAC at amino acid 515 results in an unstable protein[89] while missense mutations have been reported to affect the ability of Smad4 to form complexes with R-Smads *in vitro*[90]. Mutations found at the MH1 domain affect protein stability due to a higher susceptibility to ubiquitin-mediated degradation and they can also affect DNA binding or nuclear translocation[91,92]. Finally, there is a strong correlation between high frequency of Smad4 inactivation (either by deletions or point mutations) and distant metastasis in colorectal cancer[93–95].

WNT PATHWAY OVERVIEW AND DEREGULATION IN CRC

The WNT pathway is commonly divided into β-catenin dependent (canonical) and independent (non-canonical) signaling. For this thesis, we will focus on describing and understanding the canonical WNT signaling pathway.

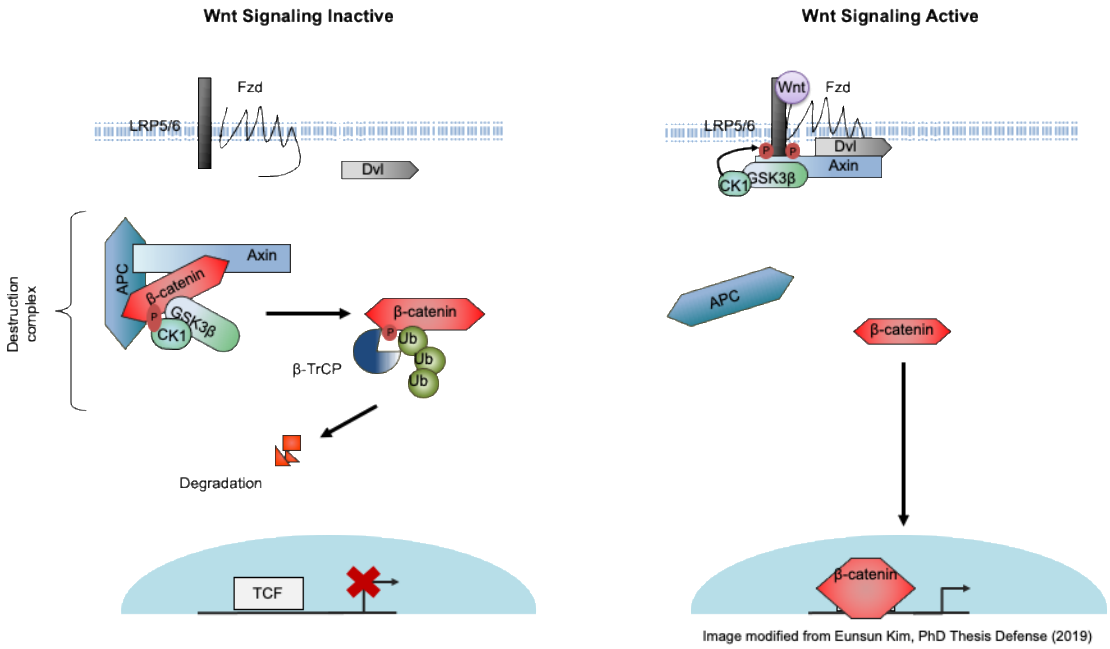

Figure 4. Canonical Wnt/β-catenin dependent signaling pathway

The canonical WNT/β-catenin signaling pathway plays different cellular roles including regulating cellular proliferation, differentiation, tissue homeostasis and repair, and cellular apoptosis[96]. In the absence of WNT ligands (WNT-inactive state), β-catenin is phosphorylated by the destruction complex which leads to its subsequent ubiquitination and proteasomal degradation. The destruction complex contains scaffold protein Axin (which contains binding domains for β-catenin), tumor suppressor gene APC and the kinases GSK3β and casein kinase 1 (CK1)[14]. When in complex, β-catenin will be phosphorylated at Ser45 by CK1, followed by phosphorylation at Ser33, Ser37 and Thr41 by GSK3β[15,97,98]. Phosphorylated β-catenin will be recognized and ubiquitinated by E3 ligase β-TrCP [99–101] and targeted for degradation. Of note, the phosphorylation sites in β-catenin are contained within the N-terminal domain, which is hotspot for mutations in cancers[102–104].

In a WNT-active state, secreted WNT ligands will bind to Frizzled (Fzd) receptors and LRP5-6 co-receptors. LRP receptors will be phosphorylated by CK1 and GSK3β, which recruits Dishevelled (Dvl) proteins to the plasma membrane, where they polymerize and are activated[14]. This will ultimately result in the dissociation of the destruction complex, which allows β-catenin stabilization and translocation into the nucleus. There, β-catenin forms an active complex with lymphoid enhancer-binding factor 1 (LEF1) and T-cell factor (TCF) proteins and a host of co-activators to drive transcription of target genes[105]. Because it has no intrinsic ability to bind to DNA, β-catenin relies on DNA-binding factors to regulate gene expression[106]. Just like with the TGF-β signaling pathway, transcription of genes driven by WNT activation vary and depends on cell type[96].

As mentioned in the beginning sections of this introduction, mutations in components of the canonical WNT signaling pathway occur in the majority of colorectal cancers with inactivating mutations in APC being most common. Mutations of APC were first identified in the germline of patients with familial adenomatous polyposis (FAP), a

hereditary condition that predisposes patients to intestinal polyps[107]. In colorectal cancer tumors, both alleles at the APC locus are affected by point mutations that can lead to stop codons or frameshifts leading to the deletion of the C-terminal half of the protein[14]. These mutations typically remove the interaction sites with Axin, leading to the inappropriate assembly of the β-catenin destruction complex[16]. β-catenin, however, is mutated in a small percentage of colorectal carcinomas (around 10%)[14,108]. The mutations affect specific serine and threonine residues which are essential for the targeted degradation of β-catenin[109].

LEF1 TRANSCRIPTION FACTOR FAMILY

Because LEF1 is a key transcription factor downstream WNT signaling, we will detail the role of LEF1 protein in WNT signaling for the remainder of this section. LEF1 belongs to the T cell factor (TCF)/LEF family of transcription factors, containing a highly conserved high mobility group (HMG) DNA-binding domain and plays a role of nuclear effects in the WNT/β-catenin pathway[110]. In the absence of nuclear β-catenin, LEF1 is bound to Groucho-related co-repressors, which recruit histone deacetylates (HDACs), resulting in the repression of WNT target genes[111,112]. Stabilization and translocation of β-catenin into the nucleus displaces Groucho complexes when binding to TCF/LEF1[113] and the active complex recruits histone-modifying coactivators such as CBP/p300[114,115].

TCF/LEFs functional roles are not just as activators, but some family members can also act as repressors. For example, while TCF1 and LEF1 are more linked to WNT target gene activation, TCF3 is generally known to be a WNT repressor[106,110]. On the other hand, TCF4 has been assigned either activating or repressor, depending on the isoform expressed. For instance, different TCF4 spliced isoforms have been studied in hepatocellular carcinoma cells. Those isoforms that contained an SxxSS motif were found

to be growth suppressive, whereas those lacking the element were WNT activating and growth promoting[116,117].

Altered LEF1 expression and function commonly occur in several cancers[118]. In many cases, increased LEF1 expression results in poor prognosis in some solid cancers including colorectal cancer[119,120], oral squamous cell carcinoma[121] and lung adenocarcinomas[122]. Given LEF1's central role as a transcription factor in the WNT/β-catenin signaling pathway, it makes for a potential target therapeutic cancer treatment. To illustrate, knockdown of LEF1 in colon cancer cells increased apoptosis *in vitro* and reduced tumor growth formation *in vivo* compared to control cells[123]. In glioblastoma cell lines, LEF1 knockdown inhibits invasion, migration, proliferation and self-renewal capacity of stem-like cells[124].

CURRENT COLORECTAL CANCER THERAPIES AND TREATMENTS

CRC is a very heterogeneous disease and evidence over the years points to the fact that molecular and genetic features of a tumor can determine both the prognosis and the response to targeted treatment. Usually patients with colon cancers up to stage II can be treated by surgical resection only. Although the cancer might have grown through all the layers of the intestine and into nearby organs, colon cancer cells have not spread to neighboring lymph nodes. For stage II cancers, it is at the doctor's discretion on whether or not to provide chemotherapy treatments to the patient[125]. Given the advances in primary and adjuvant treatments, the median overall survival time in CRC has seen much improvement. However, nearly 25% of patients with CRC present with metastases at the time of diagnosis and another 25% will subsequently develop metastasis[26,126,127].

Stage III colon cancers have spread to nearby lymph nodes, but they have not yet spread to other parts of the body. Adjuvant therapy is standard for stage III tumors, where

a combination of 5-fluorouracil plus oxaliplatin is used (where the mode of administration will depend on the protocol followed)[128]. Most people with stage IV cancer will get chemotherapies and/or targeted therapies to control the cancer. The goal of targeted therapy is to attack the cancer cells without damaging the surrounding 'normal' cells, thus leading to fewer side effects. Although each type of targeted therapy works a bit differently, they all interfere with the ability of the cancer cell to grow, divide, repair and/or communicate with other cells[129]. These therapies can include, but are not limited to, monoclonal antibodies against EGFR (cetuximab and panitumumab) and monoclonal antibodies against VEGF-A (bevacizumab)[130], which targets angiogenesis.

Despite the overwhelming evidence that both loss of TGF-β and hyperactivation of WNT pathway drive CRC, targeted therapies to the respective signaling pathways have yet to make it to the clinic. Several inhibitors/drugs that block TGF-β signaling might include small molecular inhibitors (SMI), neutralizing antibodies and others. Most current TGF-β signaling inhibitors aim to directly inhibit either TGF-β receptor kinase activity or TGF-β cytokine function[131]. Some drugs have received attention in the clinics. For example, LY33022859 is an anti-TGFβRII monoclonal antibody that inhibits receptor-mediated TGF-β-signaling activation. Although tested in a subset of patients with advanced solid tumors, it was determined that dose escalation for this drug was unsafe due to negative symptoms[132]. Fresolimumab (GC1008), a human monoclonal antibody neutralizing TGF-β-1/2/3, was tested in phase 1 with patients with malignant melanoma, also resulting in many side effects and no significant clinical benefits[133]. Generally speaking, drugs targeting receptor kinase activity can lack absolute specificity and, at certain doses, also target activin and nodal signaling pathways[131].

Theoretically, WNT-targeting approaches can be divided into three categories: 1) targeting the WNT-ligand-receptor interface, 2) regulation of β-catenin destruction

complex and 3) direct interference with β-catenin mediated transcription[134]. Vantictumab (OMP-18R5) is an antibody targeting Frizzled receptors[135] and Ipafricept is a Frizzled decoy receptor[136]. While both drugs have shown acceptable safety profiles for cancers such as breast, ovarian and pancreatic, neither is currently under investigation in CRCs. The reasoning behind, potentially, is that the majority of CRCs activate WNT signaling independent of WNT ligand. Because β-catenin is the central effector of the WNT signaling pathway, it seems the most reasonable approach to block WNT hyperactivation is by inhibiting its transcriptional response. However, complete ablation of β-catenin can be toxic to normal intestinal epithelium[137]. ICG-001 (PRI-724) is a small molecule that blocks interaction of β-catenin with its co-activator CBP[138]. Although initially developed as a potential treatment for multiple solid tumors, phase 1 trials for CRC indication are ongoing.

Although targeted therapies have been great in providing a better quality of life for many CRC patients, it is also important to understand that they have certain limitations as they only work in tumor subtype specific manner and are limited by the tumors mutational profile. Understanding the functional role of CRC mutations is a pre-requisite for development of combination therapies.

SUMMARY

The general introduction provided above serves as a base to understand the body of work presented in this thesis. Although the consequence of loss of Smad4 (absence of the protein) has been heavily studied, there is a gap in the literature in understanding a molecular function (if any) for mutant Smad4 in colorectal cancer.

CHAPTER 2: MATERIALS AND METHODS

Cell culture and treatments. Colorectal cancer cell lines SW480, SW620, and HCT116 3:6 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS) and penicillin/streptomycin (p/s) at 37 °C and 5% CO2. For the TGF-β treatment experiments, recombinant human TGF-β (R&D systems, 240-B-010) was used at a concentration of 5 or 10ng/mL at the indicated time points. For the BMP2 or BMP4 treatment experiments, recombinant human BMP2 (R&D systems, 355-BM-010) or BMP4 (R&D systems, 314-BP-010) was used at a concentration of 5 or 10ng/mL at the indicated time points. Cells were treated with recombinant Wnt3a ligand (R&D systems, 5036-WN) at 100ng/mL at indicated time points.

Transient transfections and retrovirus-mediated gene transfer. For infection of CRC cell lines with pMX-IRES constructs, retroviral particles were generated by transfection of HEK 293GP cells with Lipofectamine 3000 (Invitrogen). The virus-containing medium was collected after 8 hr or overnight and supplemented with 4 μg/ml polybrene (Sigma-Aldrich) and 1% FBS. Subsequently, the virus was filtered using 0.45 μm filter and viral supernatant was added to the target cells for 8 hours to overnight. Selection for infected cells was done with 12.5 μg/ml Blasticidin (Gemini Bio Products) for over a week.

Plasmid generation. Plasmid encoding FLAG tagged, human Smad4 (Addgene item # 14039) was first subcloned into pBlueScriptII vector and subsequently cloned into pMX-IRES-Blasticidin (Gemini Bio Products) retroviral vector. R361 Smad4 point mutations were generated using Q5 Site-directed Mutagenesis kit (NEB Cat #E0554S) with primers that carry the desired mutation. For LEF1 constructs, FLAG/MYC tagged LEF1 was

purchased from Origene (CAT #RC208663). FLAG tag was substituted with an HA tag by using Q5 Site-directed Mutagenesis kit (NEB) according to manufacturer's protocol.

Chromatin Immunoprecipitation (ChiP) qPCR. SW480 cells (2.0×10^6) were fixed by 1% formaldehyde for 10 min at room temperature and were then quenched by 125 mM glycine for 15 min at room temperature, washed with ice-cold PBS twice and centrifuged at 200g, 4°C for 5 min. The pellet was resuspended in 1 mL of cell lysis buffer (10 mM Tris-HCl, 10 mM NaCl, 0.5% NP-40, protease inhibitor) and kept at 4°C rotating for 30 min. After centrifugation, the pellet was resuspended in 1 mL of nuclear lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.0, protease inhibitor) and kept at 4°C for 60 min. Lysate was then sonicated to an average size of chromatin fragments of 0.25–1.00 kb and then frozen at -20ºC overnight. The chromatin was thawed on ice and centrifuged at 2,500g for 30 mins. Immunoprecipitation and DNA purification experiments were performed using Chip-It High Sensitivity Kit (Active Motif # 53040) according to manufacturer's protocol. The immunoprecipitated fraction was analyzed by qRT-PCR to determine the abundance of the target DNA sequence(s) relative to normal rabbit IgG control.

Primer sequences (5' to 3') used:

Table 1. Primers sequences used for ChIP qPCR

Gene Locus	Forward primer	Reverse primer
PAI-1	GCAGGACATCCGGGAGAGA	CCAATAGCCTTGGCCTGAGA

Luciferase assay. SW480, HCT116 3:6 and SW620 cells (200,000 per well) were seeded into 12-well plates and rested overnight. Cells were then transiently transfected with SBE4-Luc (Addgene #16495), pGL3-BRE-Luc (Addgene #45126), TOP-FLASH (Addgene #12456) or FOP-FLASH (Addgene #12457) reporter plasmids for 24 hrs, using renilla luciferase as internal control. For TGF-β, BMP and WNT treatments, transfected cells were stimulated with TGF-β, BMP-2/4 or Wnt3a in serum-free media for an additional 24 hrs. Firefly luciferase reporter activity was measured with a dual luciferase reporter assay kit (Promega), according to the manufacturer's protocol. Expression was calculated as the ratio of firefly luciferase units normalized to *renilla* luciferase. These experiments were independently repeated three times and each treatment consisted of triplicate samples.

Western blotting and antibodies. Whole cell protein lysates were prepared in RIPA buffer (0.15M NaCl, l% sodium deoxycholate, 0.1% sodium dodecyl sulfate, 1% Triton X-100, 1mM phenylmethylsulfonyl fluoride, 0.05M Tris-HCl, pH 7.4) containing protease and phosphatase inhibitors (Pierce Halt Inhibitor Cocktail, Thermo Scientific). Protein concentrations were estimated by Biorad colorimetric assay (BCA). Bound antibodies were detected with enhanced chemiluminescence (ECL kit, cat #) or by Odyssey Infrared Imager (LI-COR Biosciences). The following primary antibodies were used: Smad4, Smad2, Smad3, phospho-Smad2, phospho-Smad3, phospho-Smad1/5, BMPR2, Smad1, Actin, HA (Cell signaling) and FLAG (Sigma-Aldrich). For secondary antibodies, goat anti-rabbit-HRP (GE Healthcare NA934V), goat anti-mouse HRP (GE Healthcare NA931V), goat anti-mouse-680 (Licor 925-32220) donkey anti-rabbit-800 (Licor 926-32213) were used. Dilutions were used according to the recommendation of the respective manufacturers.

CRISPR/Cas9 genome-editing system. Smad4 CRISPR/Cas9 knockout and homology directed repair plasmids were purchased from Santa Cruz Biotechnologies (sc-400110 and sc-400110-HDR) and transfected into HCT116 3:6 cell lines through Lipofectamine 3000 (per manufacturer's protocol). Effective Smad4 knockdown was confirmed by western blotting.

Isolation of nuclear and cytoplasmic extract. The nuclear extraction was prepared using an NE-PER Nuclear Cytoplasmic Extraction Reagent kit (Pierce, Rockford, IL, USA) according to the manufacturer's instruction. Briefly, the treated cells were washed twice with cold PBS and centrifuged at 500 *g* for 3 min. The cell pellet was suspended in 100 µl of cytoplasmic extraction reagent I by vortexing. The suspension was incubated on ice for 10 min followed by the addition of 5.5 µl of a second cytoplasmic extraction reagent II, vortexed for 5 s, incubated on ice for 1 min and centrifuged for 5 min at 16 000 *g*. The supernatant fraction (cytoplasmic extract) was transferred to a pre-chilled tube. The insoluble pellet fraction, which contains crude nuclei, was resuspended in 50 uL of nuclear extraction reagent by vortexing during 15 s and incubated on ice for 10 min, then centrifuged for 10 min at 16 000 *g*. The resulting supernatant, constituting the nuclear extract, was used for the subsequent experiments.

Co-immunoprecipitation (co-IP). Cells stably expressing Flag-Smad4 (either wild-type or mutated) were collected and subsequently lysed in lysis buffer (50 mM Tris–HCl at pH 7.5 1M, 150 mM NaCl 5M, 1 mM EDTA, 1% Triton, 0.1% SDS, 0.5% NA-deoxycholate plus protease and phosphatase inhibitors) and incubated with anti-FLAG M2 Affinity Gel beads (Sigma-Aldrich) for 2 hrs at 4°C. The beads were washed 3 times with immunoprecipitation (IP) buffer (150 mM, NaCl, 50 mM Tris, pH 8, 1% NP-40, 0.25%

sodium deoxycholate) and bound proteins were eluted by boiling in nondenaturing sample loading buffer and loaded onto PAGE gels. For detecting binding of LEF1 and Smad4, expression constructs for LEF1-HA were transiently transfected into HCT116 3:6 cells (stably expressing Smad4) for 48 hrs. Cells were then collected, lysed and immunoprecipitation was performed as described above followed by detection through western blot.

mRNA analysis. RNA was extracted using the RNeasy Kit (Qiagen) and TRIzol (Invitrogen) and cDNA synthesis was performed using the Maxima First-Strand cDNA Synthesis kit (ThermoFisher). Quantitative PCR analysis using SYBRGreen PCR Master Mix (Thermofisher) was performed according to standard procedures.

Primer sequences (5' to 3') used:

Table 2. Primers sequences used for qPCR

Genes (all human)	Forward primer	Reverse primer
PAI-1	ACAACCCCACAGGAACAGTC	GATGAAGGCGTCTTTCCCCA
SNAI2	CATGCCTGTCATACCACAAC	GGTGTCAGATGGAGGAGGG
PRSS8	CACCTTCTCCCGCTACATCC	AGGAGGCTCACTGAGGGG
NLRP2	GCAAAGGATGAAGTCAGAGAAGC	TTTTGAAGCGCTCCAGCATT
ACTIN	AGCATCCCCCAAAGTTCAC	AAGGGACTTCCTGTAACAACG

RNA-seq. Total RNA was extracted from SW480 cells using RNeasy Mini Kit (QIAGEN, #74104) and polyA+ transcripts were isolated with NEB Next Poly(A) mRNA Magnetic Isolation Module (NEB, #7490). RNA-Seq libraries were prepared with NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, E7760S). Three biological replicates

were sequenced on a NextSeq 500 (Illumina) at a depth of at least $2x10^7$ reads each. Reads were mapped and analyzed with a bioinformatic pipeline based on GSNAP, featureCounts, and the R packages limma and edgeR. We used human genome version GRCh38. GO analyses were performed using version 6.8 of the DAVID web server. GSEA analyses were performed using pre-ranked GSEA using a weighted scoring.

Statistics and reproducibility. All statistical analyses were carried out using Graph Pad Prism (version 7) by unpaired student's t-test for two group comparisons, Yate's continuity corrected chi-square test for contingency table analysis, and Gehan-Breslow-Wilcoxon test for survival curves. Error bars represent s.e.m. ± SD and statistical significance were defined as $P < 0.05$.

CHAPTER 3: SMAD4 WITH R361 HOTSPOT MUTATIONS RETAINS THE ABILITY TO BIND TO LEF1 AND BOOSTS WNT SIGNALING IN COLORECTAL CANCER

INTRODUCTION

In normal and premalignant cells, TGF-β can exert tumor-suppressive effects by inhibiting cell proliferation, stimulating differentiation and inducing apoptosis[139]. However, tumor cells can evade TGF-β tumor suppression through inactivation of either the receptors or downstream effectors, which disables the entire signaling cascade. For example, TGF-β RII is frequently mutated in colon carcinoma cells from patients with microsatellite unstable (MSI) CRC, a phenomenon defined by faulty DNA mismatched repair machinery[5]. Inactivating mutations in Smads, specifically Smad4, have been found in various types of cancers such as lung and pancreatic carcinomas[83,87,140–142] and germ-line mutations are common in juvenile polyposis, a disease which predisposes individuals to gastrointestinal malignancies[86]. Smad4 alterations are also found in 10-35% of CRC tumors and tend to appear late in the adenoma-to-carcinoma progression[143,144]. Complete loss of Smad4 in colorectal cancer patients can be due to deep deletions or frameshift mutations. We and other have shown that it is frequently associated with increased angiogenesis[81,145], lymph node metastasis, advanced disease, and poor prognosis[84]. In addition, many patients acquire missense mutations, which cluster in the MH2 domain of the protein. While mutations in the MH2 domain mainly affect residues close to protein interface involved in hetero-oligomerization of Smad4 with R-Smads which is required for transcriptional activation[53,55], mutations in the MH1 domain have been shown to alter protein stability and binding to the DNA[49,146]. Thus, Smad4 missense mutations are thought to be loss-of-function. However, there is a considerable gap in our understanding

of their role in CRC progression and whether these mutant forms of Smad4 retain or acquire any functions.

In this results section, we report that R361 Smad4 variants can function independently from TGF-β-signaling and positively regulate WNT signaling, a pathway often hyperactivated in CRC[16]. We demonstrate that mutant Smad4 binds to lymphoid enhancer binding factor-1 (LEF1) protein and facilitates transcriptional activation of WNT signaling in CRC cells. Overall, we establish a novel function for mutant Smad4 proteins in the progression of colorectal cancer.

RESULTS

SMAD4 is frequently mutated in colorectal cancer

To understand the clinical relevance of SMAD4 mutations in CRC, we analyzed SMAD4 mutation data in recent Memorial Sloan Kettering Cancer Center (MSKCC) study, which generated cancer gene panel data from 1,134 colorectal adenocarcinomas from patients with both metastatic and early-stage CRC. While all the TGF-β signaling components exhibit some genetic alteration, SMAD4 was the most frequently altered gene from the TGF-β signaling pathway (Figure 5A)[21]. Of all SMAD4 mutations and copy number alterations, missense mutations classified as putative drivers were the most abundant. Although they could be found across the MH2 domain, we confined subsequent analyses to amino acid substitutions in the R361 residue, as they were found at a much higher frequency than other missense mutations in this data set and other CRC studies[143,147] (Figure 5B).

To study R361 mutations in CRC cells, we generated retroviral constructs expressing either Smad4 wildtype (WT) or point mutations at the MH2 domain of Smad4, specifically at the arginine 361 residue (R361) and bearing FLAG tag at the N-terminus.

We transduced those constructs into cells lines that lack detectable endogenous Smad4 protein, SW480 and SW620[88], but have otherwise intact TGF-β signaling pathway components[148] (Figure 6A). After transduction with the retroviral constructs, we were able to detect robust protein expression by Western blotting using the α-Smad4 antibody (Figure 6B). To study the Smad4 mutations in additional TGFβ-responsive cell lines, we used an HCT116 engineered derivative HCT116 3:6, which have intact TGF-β signaling[149]. To render them Smad4-deficient, we used Santa Cruz CRISPR-Cas9 knockout (KO) and homology directed repair (HDR) plasmid according to manufacturer's protocol. Edited HCT116 3:6 cells were cloned using flow cytometry and a single cell clone with no trace of Smad4 expression was identified using Western blotting (Figure 6C & 6D). When comparing the level of expression between parental HCT116 3:6 cells (which express endogenous Smad4 protein) and our retroviral constructs, we observe that HCT116 3:6 cell lines transduced with Smad4-retroviral constructs express ~2-3X as much Smad4 protein as those with endogenous Smad4 (Figure 6E). Thus, although levels of retrovirally expressed Smad4 didn't exactly match those seen in CRC, they were not grossly elevated either. Finally, the effects of R361-mutation on the turnover of Smad4 was examined in SW480 cells using cycloheximide treatment (25ug/mL for 0, 6 and 12 hrs). It was observed that both Smad4-WT and Smad4-R361H/C protein start degrading between 6-12 hrs (Figure 6F), suggesting that mutations at the R361 residue did not affect the half-life of Smad4 protein.

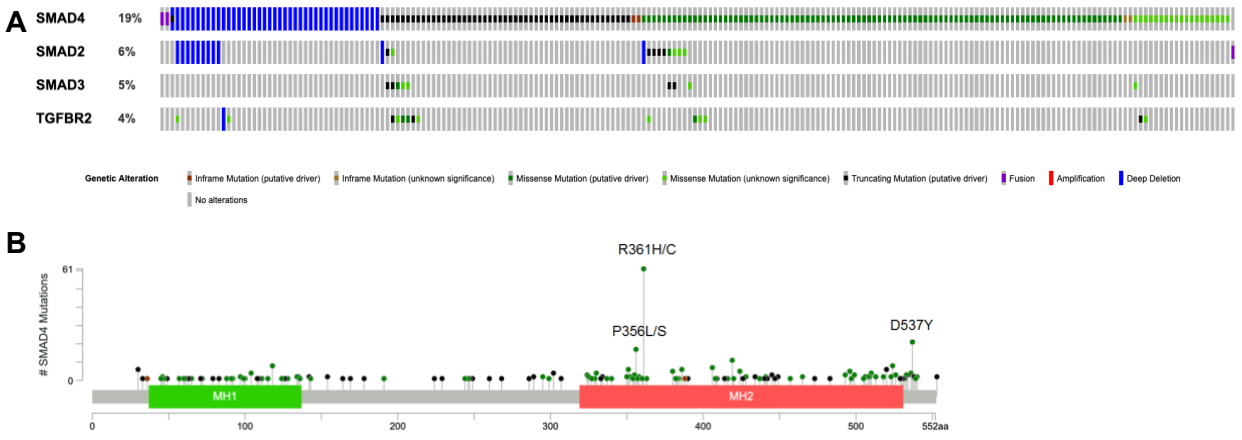

Figure 5. Smad4 is frequently mutated in CRC. (A) MH2 hotspot mutations of SMAD4 are marked on the lollipop plot downloaded from the MSKCC (2018) study on cBioportal. **(B)** The prevalence and spectrum of SMAD4 mutations in MSKCC study.

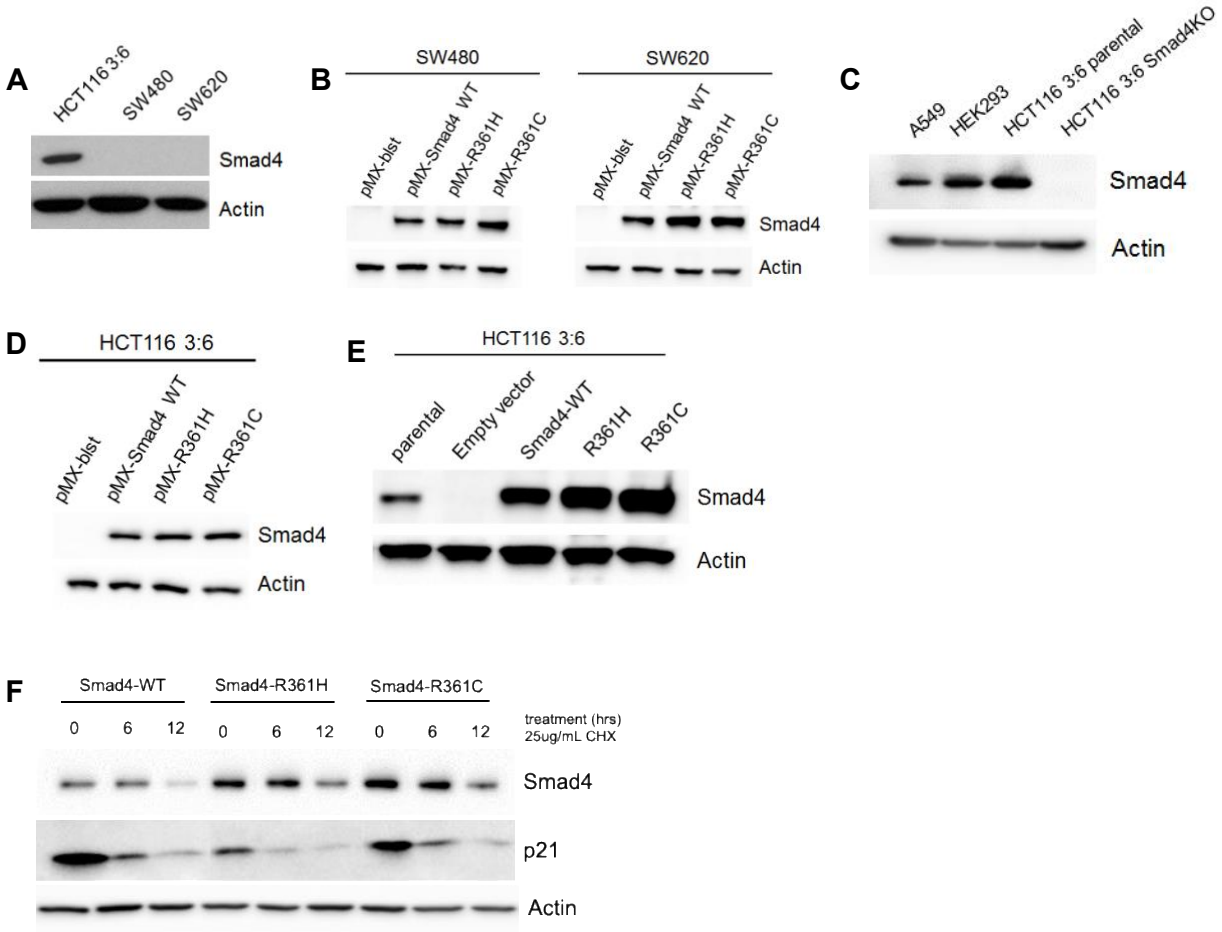

Figure 6. Generation of Smad4 CRC mutant cell lines. (A) Western blot analysis of cells lacking endogenous expression of Smad4, using HCT116 3:6 as positive control. **(B)** Western blotting confirming retroviral expression of Smad4 in SW480 and SW620 cell lines. **(C)** Western blot comparing levels of Smad4 protein in A549, HEK293, HCT116 3:6 and HCT116 3:6 Smad4 KO cell lines. Actin was used as loading control. **(D)** Western blotting confirming retroviral expression of Smad4 in HCT116 3:6 cells. **(E)** Western blot comparing the levels of endogenous versus retroviral transduced Smad4 protein. Actin was used as a loading control. **(F)** Determination of Smad4 half-life. SW480 cells were treated with 25ug/mL of cycloheximide for either 0, 6 or 12 hrs, followed by western blot with the indicated antibodies. p21 was used as a positive control and actin was used as a loading control.

SMAD4 R361 missense mutations prevalent in colorectal cancer result in loss of binding to phosphor-Smad2/3

Because Smad4 is an important factor in TGF-β signaling, we asked whether there were major differences in TGF-β upstream signaling pathway. Stimulation of these two cell lines for 1 hour with 5ng/mL of TGF-β resulted in phosphorylation of receptor regulated Smads (i.e., Smad2 and Smad3, Figure 7A) regardless of Smad4 status. This was expected since phosphorylation of Smad2/3 is an initial step in the TGF-β signaling cascade and upstream of Smad4 signaling. However, R361 is in the loop helix region of the Smad4 MH2 domain and is an important residue in forming a Smad4 heterocomplex with R-Smad[38]. To test whether mutations at the R361 residue are still able to bind to phosphorylated Smad2 and Smad3, we performed immunoprecipitation analysis on SW480 cells treated with either vehicle control or soluble TGF-β. While wild-type Smad4 was able to bind to both R-Smads, both mutations R361H and R361C abolished interaction with R-Smads in CRC cells (Figure 7B). This suggests that this missense mutation affects heterodimerization between Smad4 and Smad2/3 but does not affect overall phosphorylation of Smad2/3.

In a basal state, Smad4 is distributed throughout the cell and may undergo continuous shuttling between the cytoplasm and nucleus – presumably due to the combination of an active nuclear localization signal (NLS) in its MH1 domain and its nuclear export signal (NES) in its linker region. During active TGF-β signaling, binding of R-Smads to Smad4 is thought to mask the NES in Smad4, thus allowing the complex to accumulate in the nucleus[49]. Because R361 mutations are unable to associate with R-Smads, we decided to investigate whether or not Smad4 mutations can concentrate into the nucleus following TGF-β treatment. Following treatment with soluble TGF-β ligand, only wild-type Smad4 was able to accumulate in the nucleus while Smad4-R361H and Smad4-R361C mutants could not. (Fig. 7C, 7D quantification). These results suggest that

Smad4-R361 mutations are unable to bind to R-Smads and cannot concentrate in the nucleus in the presence of active TGF-β signaling.

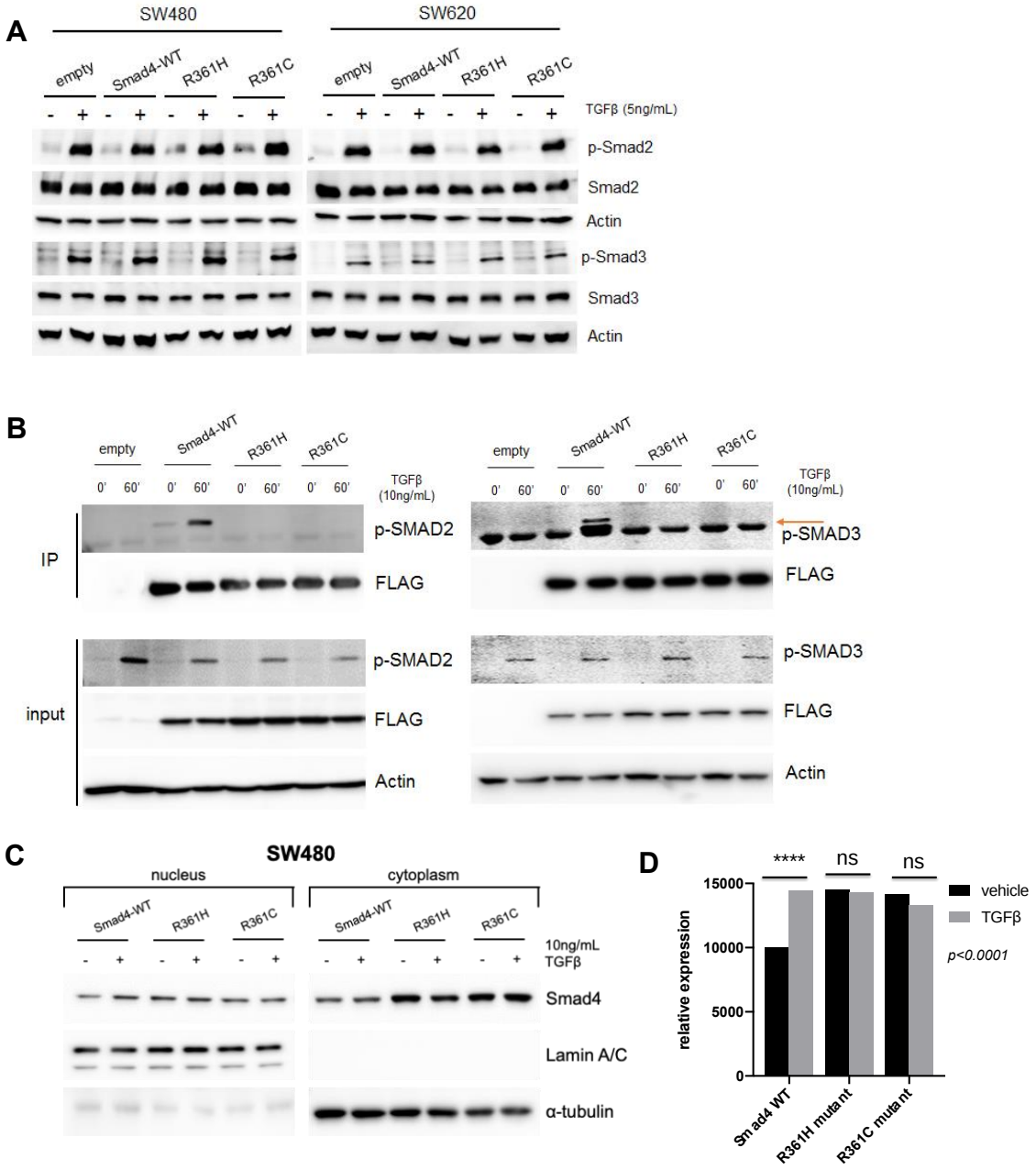

Figure 7. Smad4-R361 mutations cannot bind p-Smad2/3 in CRC cells in response to TGFβ. (A) SW480 and SW620 cell lines treated with either vehicle control or 5ng/mL for an hour. Western blotting was performed to detect phosphorylation of Smad2 and Smad3, and total levels of Smad2 and Smad3 as indicated. Actin was used as loading control. **(B)** SW480 cells were treated at 0' or 60' with soluble TGF-β followed by immunoprecipitation with FLAG conjugated beads. Input/Whole Cell Lysate (WCL) was used as a control. **(C)** SW480 cells were treated at 0' or 60' with soluble TGFβ followed by a nuclear cytoplasmic fractionation. Lamin A/C was used as nuclear marker and alpha tubulin was used as cytoplasmic marker. **(D)** Quantification of nuclear Smad4 in SW480 experiment.

Smad4 mutants do not support canonical TGF-β-induced transcriptional activity

Activated Smad complexes accumulate in the nucleus, where they can bind directly or indirectly to DNA and regulate gene expression. Because Smad4 is a critical effector of TGF-β signaling, we tested whether R361 mutations in Smad4 have altered transcriptional activity. Smad4 wild-type and Smad4 mutant stably expressing cells lines were transiently transfected with a luciferase vector containing 4 copies of the Smad binding element (SBE), a sequence that allows for Smad-DNA binding[43]. As expected, Smad4-WT was able to support downstream TGF-β transcriptional activity in the presence of soluble TGF-β in all three of our CRC cells lines (SW480, SW620 and HCT116 3:6). However, both Smad4-R361 mutations were unable to support canonical TGF-β transcriptional activity and behaved similarly to the empty-vector cells (Figure 8A). There has been uncertainty as to whether mutations in Smad4 could exert a dominant-negative effect over Smad4-WT[62,89]. We therefore tested the ability of mutant Smad4 to interfere with Smad4- mediated TGF-β responsive transcription. Due to ease of transfection, we used the HCT116 3:6 Smad4 knockout (KO) clone cell line to test whether or not mutant Smad4 acted in a dominant negative manner over wild-type Smad4. Transfection of increasing amounts of either Smad4-R361H or -R361C had no effect on induction of SBE reporter gene by TGF-β (Figure 8B) arguing against the dominant negative mechanism. To further confirm the decrease in transcriptional activity of Smad4-R361 mutants, we performed quantitative analysis of Smad4-targeted gene expression by real-time quantitative RT-PCR (qRT-PCR). The expression of PAI-1 and SNAI2, both well-validated direct Smad4 target genes, was compared in SW480 cells expressing different Smad4 constructs and the empty vector. Consistent with the reporter assay, in TGF-β-treated SW480 cells, only Smad4-WT was able to induce the expression of these Smad4 target genes by more than 2-fold, whereas this induction was not seen in cells overexpressing Smad4- R361 mutations (Figure 8C).

Smad4 binds directly to the promoter region of PAI-1[150]. To provide further evidence that mutant Smad4 is unable to regulate PAI-1 gene expression in the presence of TGF-β, we performed chromatin immunoprecipitation (ChIP) assay followed by qRT-PCR. As expected, stimulation of cells with TGF-β induced the recruitment of Smad4-WT, but not Smad4-R361 mutant protein to the promoter region of PAI-1 (Figure 8D). Taken together, these data indicated that Smad4-R361 mutations completely disengage from the TGF-β pathway and neither activate nor repress canonical TGF-β signaling, even when overexpressed.

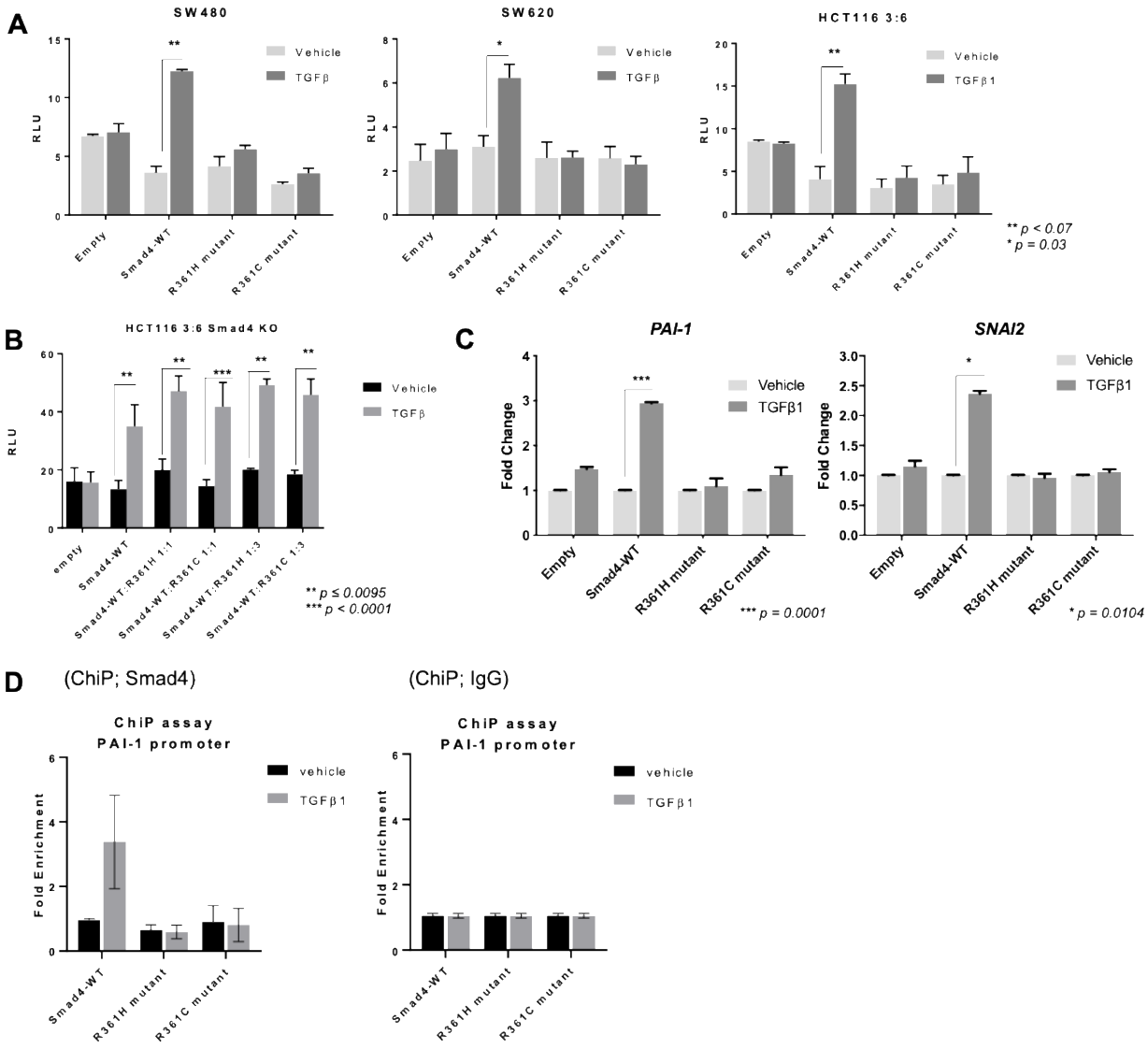

Figure 8. Stably expressed Smad4 mutants do not support canonical TGFβ-induced transcriptional activity. (A) SW480, SW620 and HCT116 3:6 cells transfected with the SBE-Luciferase construct, later treated with either vehicle control or soluble TGFβ ligand. pRL-TK plasmid was used as an internal control. **(B)** HCT116 3:6 cells were transiently transfected SBE-Luciferase construct, with either a 1:1 or a 1:3 ratio of Smad4-WT:Smad4-R361 mutant plasmid for 24 hours, followed by treatment with either vehicle control or soluble TGFβ for another 24 hours. Cells were then lysed, and luciferase activity in lysates was determined. Renilla plasmid was used as an internal control. **(C)** Induction of PAI-1 and SNAI2, TGFβ target genes, were measured after TGFβ treatment by qRT-PCR in SW480 cells. Actin was used as internal control. **(D)** SW480 cells were plated and treated with 10ng/mL of soluble TGFβ for an hour. Chromatin fraction was prepared from harvested cells to analyze binding of Smad4 to indicated promoter using ChIP-qPCR with anti-Smad4 or anti-IgG.

Smad4 mutants do not support canonical BMP-induced transcriptional activity

Although our studies so far have mainly focused on the TGF-β signaling pathway, it is important to reiterate that Smad4 can mediate gene responses in a variety of other pathways. One of the other well studied TGF-β signaling pathways is that of Bone Morphogenic Protein (BMP). BMP ligands also bind to type II and type I serine-threonine kinase receptors and mediate their signaling through both Smad and non-Smad signaling pathways. As mentioned in Chapter 1 of this thesis: of the eight Smad proteins identified in mammals Smad1, Smad5 and Smad8 are the R-Smads activated by BMP type I receptors[151]. While BMPs can be divided into several subgroups, BMP-2/4 group has been most recently identified to play a role in colorectal cancer[152].

Due to the fact Smad4 (co-Smad) is shared by both BMP and TGF-β signaling pathways, we also wanted to understand whether or not mutations in Smad4 affects BMP signaling. Because phosphorylation of R-Smads is the initial step in activation, we decided to look at the effect of BMP upstream signaling pathway. Stimulation of SW480 for 1 hour with 50ng/mL of either BMP2 or BMP4 resulted in phosphorylation of receptor regulated Smad1/5 in cell lines expressing wild-type Smad4. Interestingly, we see that cells containing no Smad4 or Smad4-R361 mutation have reduced phosphorylation of R-Smads (Fig 9A), an event that we did not see when we treated cells with TGF-β and analyzed phosphorylation of Smad2/3 (Fig 7A). This event, however, would need to be confirmed in the other CRC cell lines used in our study.

Similarly, in our SBE-luciferase assays, Smad4-WT was able to support downstream BMP transcriptional activity in the presence of soluble BMP2 and BMP4 in SW480 and HCT116 3:6 cell lines. However, both Smad4-R361 mutations were unable to support canonical BMP transcriptional activity and behaved similarly to the empty-vector cells (Fig 9B). Finally, to test whether Smad4-R361H is still able to p-Smad1/5, we performed immunoprecipitation analysis on SW480 cells treated with either vehicle control

or soluble BMP2 and BMP4. Similar to our p-Smad2/3 assay, only wild-type Smad4 was able to bind to p-Smad1/5 as Smad4-R361H completely abolished that interaction (Fig 9C). Overall, mutating R361 residue affects heterodimerization of Smad1/5 and downstream BMP signaling.

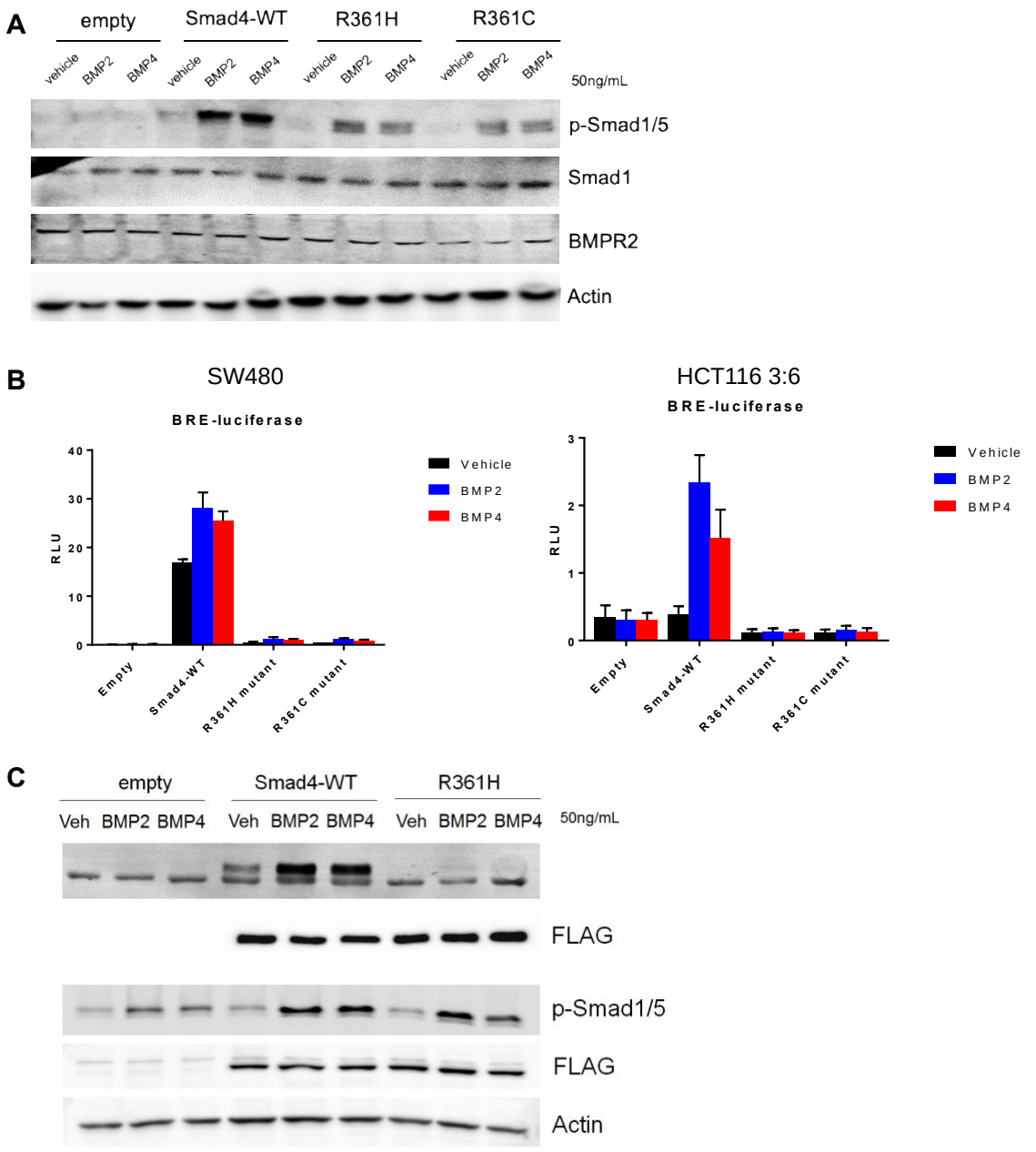

Figure 9. Smad4-361 mutations result in loss of downstream BMP signaling (A) SW480 cell lines treated with either vehicle control or 50ng/mL of BMP2 or BMP4 ligand for an hour. Western blotting was performed to detect phosphorylation of Smad1/5, and total levels of Smad1 and BMPR2 as indicated. Actin was used as loading control. **(B)** SW480 and HCT116 3:6 cells transfected with the BRE-Luciferase construct, later treated with either vehicle control or 50ng/mL of BMP2 or BMP4 ligand. pRL-TK plasmid was used as an internal control. **(C)** SW480 empty, Smad4-WT and R361H expressing cells were treated with either vehicle, BMP2 or BMP4 ligand followed by immunoprecipitation with FLAG conjugated beads. Input/Whole Cell Lysate (WCL) was used as a control.

R361H confers the gene signature associated with LEF1 protein overexpression.

Although with respect to canonical TGF-β targets Smad4-R361 mutations appear to be loss-of-function, we reasoned that a hotspot mutation is more consistent with a gain-of-function mechanism. To uncover such a mechanism, we performed RNA-Seq analysis on SMAD4-transduced SW480 cells treated with soluble TGF-β for 24 hr (Figure 10A). Principal component analysis (PCA) performed on all expression datasets revealed that as expected, Smad4-WT is the only group strongly affected by the treatment with soluble TGF-β (Figure 10B). However, unexpectedly, the SMAD4 mutant samples separated very strongly from the SMAD4-null samples, with or without TGF-β treatment. This separation was the first experimental evidence of nonequivalence of SMAD4-null and missense mutations. To understand the specific differences between CRC cells expressing no Smad4 and Smad4-R361 mutations, we further analyzed our RNA-Seq data set by gene set enrichment analysis (GSEA). We discovered that Smad4 R361H-associated genes are most significantly enriched in the members of the following datasets: genes down or up-regulated in DLD1 cells (colon carcinoma) over-expressing LEF1 (Figure 10C). This enrichment was strongly driven by the cluster of genes that were expressed at high levels in SMAD4-null cell but downregulated in SMAD4-WT and especially in SMAD4-MUT cells (Figure 10D) such as known colon cancer metastasis suppressor PRSS8[153], putative CRC tumor suppressor RAB25[154], and an NF-kB inhibitor NLRP2[155]. To confirm that a transcriptome associated with LEF1-levels was indeed affected, we reproduced RNA-seq results by analyzing single genes by qRT-PCR. Indeed, we saw that transcription of genes included in the list: NLRP2 and PRSS8 showed downregulation in cells containing both Smad4 R361H and R361C mutant compared to cells expressing no Smad4. Of note, RAB25 was significantly downregulated in cells expressing Smad4-R361C, but not Smad4-R61H, when compared to empty vector cells (Figure 10E).

The RNA-seq data showed that LEF1-associated gene signature is affected in cells expressing Smad4 mutant vs empty vector cells. To study this effect at the molecular level, we first determined that the level of LEF1 protein was unchanged between empty vector, Smad4-WT and Smad4-R361 CRC cells (Figure 11A). Previous studies have demonstrated that Smad4 WT and LEF1 can bind each other in neuroblastoma and fibroblast-like cells[156,157]. To investigate whether or not Smad4-R361 mutants can bind LEF1 protein in CRC cells, we transiently transfected Smad4-transduced HCT116 3:6 cultures with an HA-tagged LEF1 construct followed by immunoprecipitation. We observed that both Smad4-WT and Smad4 R361H/C can bind LEF1 protein in CRC cells (Figure 11B), but the consequences of this interaction remained to be determined.

It's been previously shown amino acids 511-552 are crucial for interaction of SMAD4 with LEF1[157]. However, when we tried to generate a Smad4Δ510-552 construct, we are unable achieve Smad4 protein expression as examined through western blot analysis (data not shown). To determine the domain essential for Smad4-LEF1 interaction, we generated two Smad4 deletion constructs: Smad4Δ535-552 and Smad4Δ543-535 (Figure 11C). We first transfected these constructs into HEK cells to validate their expression (Figure 11D) via western blot, and then tested whether or not they could retain binding to LEF1. Again, we transiently transfected Smad4-transduced HCT116 3:6 cultures with an HA-tagged LEF1 construct followed by immunoprecipitation. We observed that while Smad4-WT can bind to LEF1 protein, that binding is significantly reduced in both Smad4Δ535-552 and Smad4Δ543-535 samples (Figure 11E and 11F, quantification). This led us to believe that Smad4 amino acids 535-552 might be essential for binding to LEF1.

As LEF1 is a key downstream effector of the WNT pathway, we evaluated WNT signaling status in our CRC cell lines. Specifically, we employed the TOP-Flash/FOP-Flash luciferase reporter that contains 7 copies of TCF/LEF binding sites. When incubated

in basal media, R361 Smad4 mutants were able to induce downstream WNT signaling at higher levels compared to no-Smad4 and wild-type Smad4 in both SW480 and HCT116 3:6 cells (Figure 12A). Since SW480 are not responsive to Wnt3a (data not shown), we continued our studies with HCT116 3:6 cells. When HCT116 3:6 Smad4 cells were treated with soluble Wnt3a ligand, we observe an increase in WNT signaling compared to vehicle control. In addition, there was significant up-regulation of WNT signaling by R361 mutants when compared to cells expressing no Smad4 and wild-type Smad4 (Figure 12B). We conclude that these two Smad4 mutations, which are associated with the loss of TGF-β response in human cancer cells, still retain a TGF-β-independent function of Smad4, activation of downstream WNT signaling.

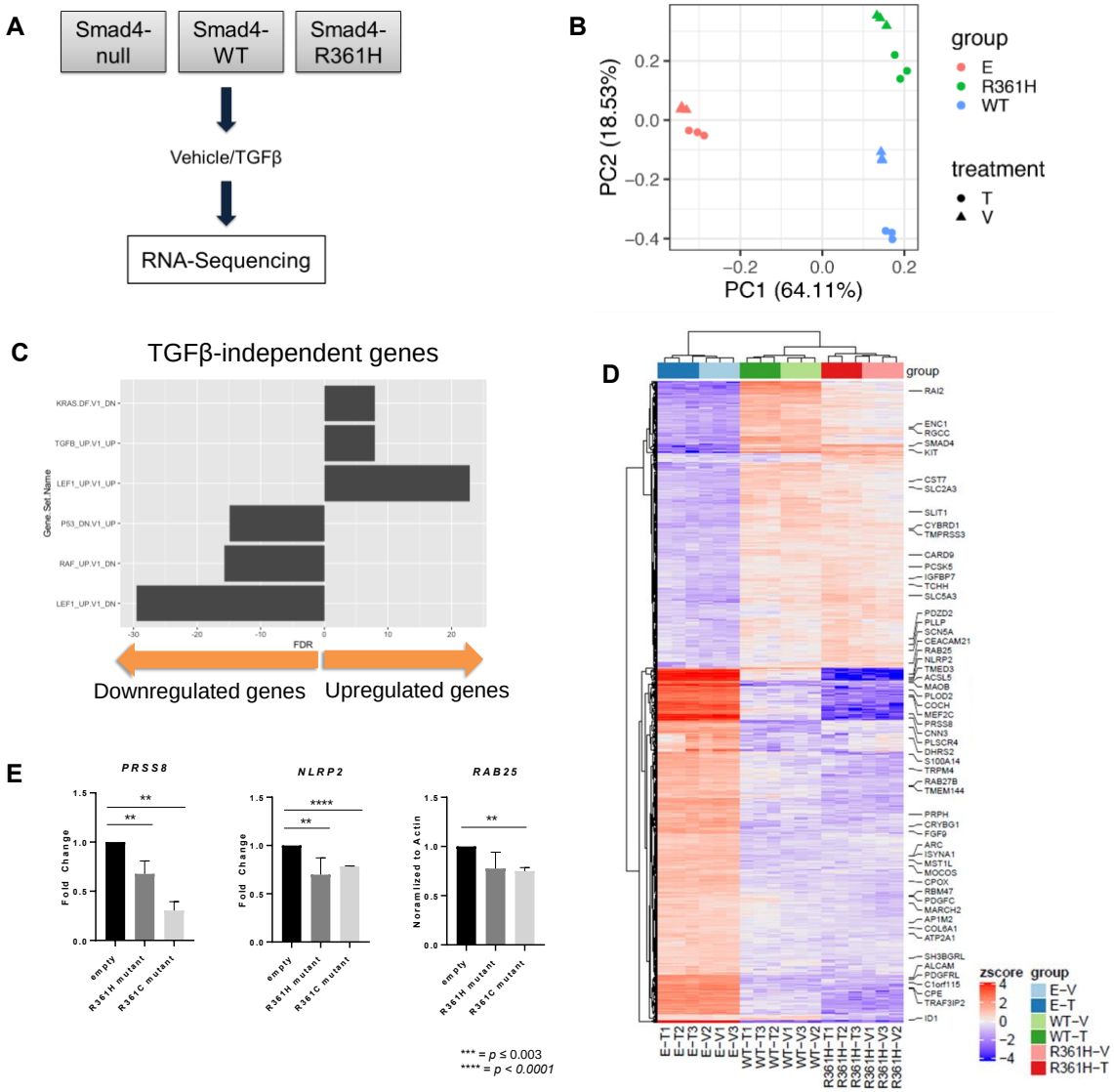

Figure 10. R361H has the gene signature associated with upregulation of TCF/LEF1 protein complex. (A) Schematic representation of our RNA-Seq workflow. **(B)** Principal component analysis (PCA) for different treatment groups as indicated. **(C)** GSEA summary of pathways associated with genes differentially regulated in cells expressing Smad4-R361H vs empty vector expressing cells. **(D)** Heat map representing the gene expression in CRC cells expressing Smad4-R361H mutant, Smad4- WT and empty vector **(E)** qRT-PCR validation on NLRP2, PRSS8 and RAB25 transcript on SW480 cells. Actin was used as endogenous control and empty vector cells were used to normalize expression of all three genes.

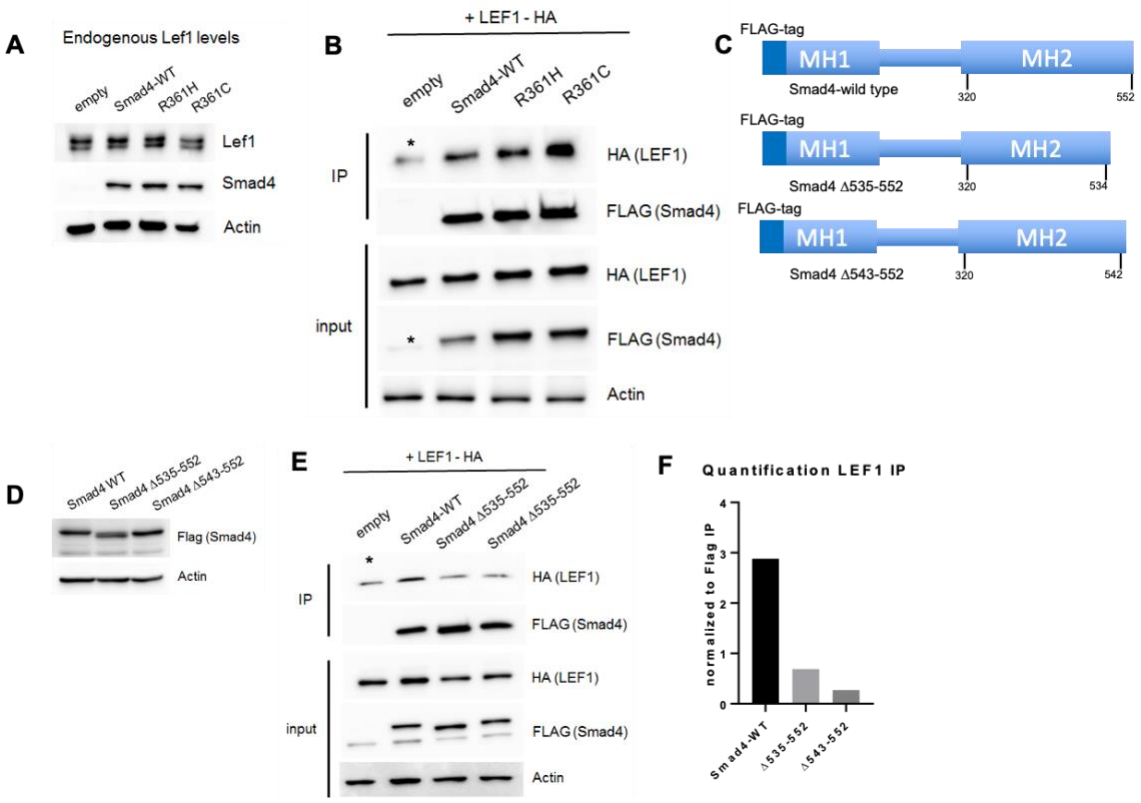

Figure 11. Smad4-binds to LEF1. (A) Western blotting showing the levels of Lef1 protein in SW480 cell lines. **(B)** HCT116 3:6 cells were transfected with Lef1-HA constructs for 48 hours followed by immunoprecipitation with FLAG-conjugated beads. Input/Whole Cell Lysate (WCL) was used as a control. Asterisks represent unspecific band in the empty vector lane, likely due to antibody background. Immunoprecipitation was repeated three times with the same results. **(C)** Diagram of Smad4 mutation constructs generated to test which domains are necessary to bind to LEF1. **(D)** HEK cells were transiently transfected with Smad4-WT, ΔSmad4 535-552 and ΔSmad4 543-552 to confirm expression of protein. Actin was used as loading control. **(E)** IP and western blot assays showing the interaction between Smad4 and LEF1 in HCT116 3:5 cells transfected with different FLAG-tagged Smad4 deletion constructs and HA-tagged LEF1 construct. Asterisks represent unspecific band in the empty vector lane, likely due to antibody background. **(F)** Quantification of blots from panel E.

WNT activation and SMAD4 missense mutations in primary CRCs

We then asked whether there is an epistatic relationship between Smad4 missense mutations and WNT activation in CRC patients. To answer this question, we re-analyzed the MSKCC cancer panel study. In microsatellite stable CRC samples, SMAD4 has a 15% frequency of mutation while CTNNB1 (β-catenin) has a lower mutation frequency of 4% (Figure 12C), likely due to APC already being mutated in many CRC patients[14]. We reasoned that if SMAD4 missense mutations aid WNT signaling, activating missense mutation in CTNNB1 will be under even less selective pressure to occur. This would be in contrast to SMAD4 truncating mutations, which do not affect WNT signaling. To compare and contrast the co-occurrence frequency of CTNNB1 mutations in CRCs alongside either SMAD4 missense or loss-of-function nonsense mutations[89,158], we utilized the chi-square test with Yate's correction. Indeed, we observed highly statistically significant enrichment for CTNNB1 mutations in CRC with nonsense mutations, at the expense of tumors with SMAD4 missense mutations (Figure 12D).

Finally, to understand the contribution of R361-mutations to clinical outcomes in CRC patients, we analyzed median overall survival (OS) in the MSKCC cohort. When compared with patients with no SMAD4 alterations, patients with SMAD4 missense mutations had a shorter OS survival of 70.03 vs 40.5 months ($p = 0.0186$). This difference was not be observed in patients that had no alteration for SMAD4 versus patients with SMAD4 truncating nonsense mutations (OS survival of 70.03 vs 68.3 months respectively, $p = 0.9076$) (Figure 12E, left and right). These data demonstrate that CRC patients with SMAD4 missense mutations do worse overall than those with unaltered SMAD4 and truncating mutations in SMAD4, attesting to their gain-of-function properties.

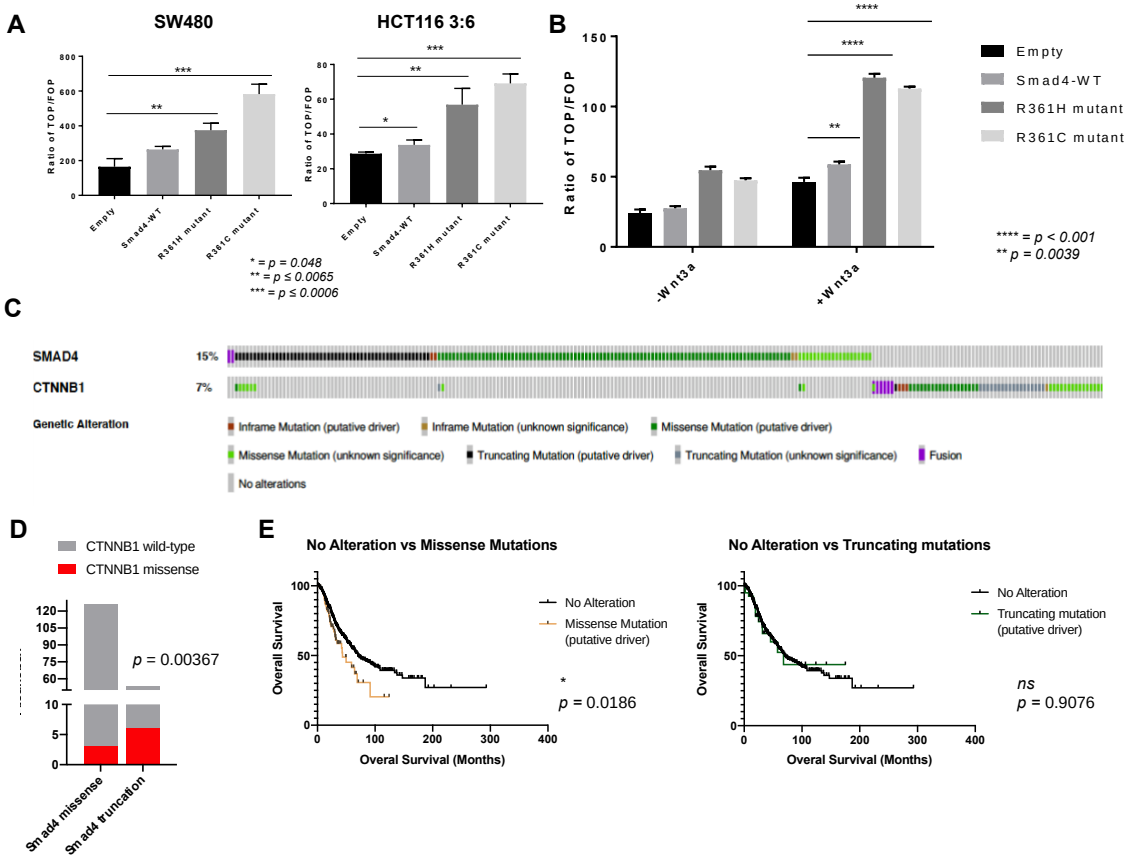

Figure 12. Smad4 R361 mutants boost Wnt signaling. (A) SW480 and HCT116 3:6 cells were transfected with either a TOP-flash or FOP-flash construct and collected after 48hrs. pRL-TK plasmid was used as an internal control. **(B)** HCT116 3:6 cells were transfected with either a TOP-flash or FOP-flash construct, later treated with either vehicle control or Wnt3a ligand. pRL-TK plasmid was used as an internal control. **(C)** Oncoprint of SMAD4 and CTNNB1 mutations of MSS CRC patient samples, which were then analyzed using the cBio Portal for Cancer Genomics platform. **(D)** CRC patients with SMAD4 missense mutations containing CTNNB1 missense mutations (3) or CTNNB1 wild-type (123) or SMAD4 truncating mutations containing CTNNB1 missense mutations (6) or CTNNB1 wild-type (48). Statistical significance using Yate's continuity corrected chi-square test: p = 0.00367. **(E)** Comparison of survival curves in patients with no SMAD4 alterations and SMAD4 missense mutations (left) and no SMAD4 alterations and SMAD4 truncating mutations (right). Statistical significance per log-rank (Mantel-Cox) test: p = 0.0186 and p = 0.9076 respectively.

DISCUSSION

Defects in Smad4 play a significant role in the malignant progression of tumors and are frequently altered in colon and pancreatic carcinomas[21,83]. About 10-30% of CRC patients harbor some loss of Smad4, either by deep deletion or nonsense mutations. Interestingly, missense mutations in Smad4 also have been identified in variety of cancers including CRC. In these tumors, Smad4 mutations appear frequently at the MH2 domain. By analyzing both primary tumors and immortalized cells lines, two of the most frequent Smad4 missense mutations have been identified at position 361 which results in a substitution from arginine to either histidine or cysteine[55]. These genetic hits are recurrently detected in Smad4, consistent with the idea that this gene acts as a tumor suppressor. On the other hand, their clustering in distinct hot spots argues that Smad4 mutants, similar to mutant have an oncogenic function; however, this hypothesis has not been previously tested.

In the present study, we investigated the functional roles of the missense mutations R361H and R361C in the Smad4 MH2 domain that naturally occur in human colorectal cancer patients[55]. This was achieved by the means of retroviral transduction into cell lines that lack endogenous Smad4 protein. Smad4 missense mutations had been previously mapped onto the crystal structures of Smad heterodimers, specifically onto the defined protein loop that is directly involved in binding to the R-Smads. Our co-immunoprecipitation experiments indicate that point mutations in Smad4 MH2 domain indeed disrupt binding to endogenous p-Smad2, p-Smad3 and p-Smad1/5. We also showed in transfection assays that Smad4 missense mutations cannot support transcription from reporters driven by Smad3–Smad4 (CAGA12- luciferase) complexes. In addition, Smad4-R361 mutations could also not support downstream BMP signaling when we used a BMP-specific reporter. Although it seems that these mutations also cannot activate downstream BMP signaling, more experiments would be needed to solidify

this claim. For example, apart from looking at BMP-signaling using a reporter assay, we could confirm loss of canonical BMP target genes such as *Id1* and *Id2*[153,154] via qRT-PCR. In addition, we could also determine if Smad4 can bind to the promoter of BMP target genes through Chromatin Immunoprecipitation (ChiP) as we did with TGF-β/SMAD4 target, PAI-1.

Admittedly, some tumor suppressor genes can have dominant negative effects. For example, mutant p53 protein can bind to its wild type counterpart encoded by the unaltered allele and sequester it in non-functional complexes[159]. However, when we co-expressed mutant and wild- type Smad4, we observed no dominant negative effects of the former. Taken at face value, these experiments would suggest that Smad4 variants are loss-of-function. This mechanism is supported by several lines of genetic evidence, including a frequent loss of heterozygosity (LOH) at chromosome locus 18q where SMAD4 is located; such loss is associated with a poor prognosis for CRC patients[160,161]. Yet the very high prevalence of R361 SMAD4 mutations in CRC patients, particularly in those with distant metastasis vs. locally advanced tumor[93] indicated that additional molecular mechanisms could be at play.

The idea that Smad4 can lose canonical TGF-β functionally but maintains some TGF-β-independent function had been proposed before. For example, it was suggested previously that Smad4 cooperated with LEF1 to increase c-myc expression in the absence of TGF-β signaling[162]; however, that study did not investigate the effect of this interaction on WNT signaling. Furthermore, to our knowledge, the ability of mutant Smad4 to bind to LEF1 and the consequence of that binding has not been previously studied. Here, we propose that Smad4-R361-mutant loses its tumor suppressive arm and retains binding to LEF1, resulting in enhanced WNT signaling. We do not know for the fact why there is more WNT signaling in Smad4-R361 CRC expressing cells compared to those expressing Smad4-WT. It is possible that the inability of Smad4 R361 mutant to bind R-Smads

following TGF-β stimulation increases the pool of available protein to bind to LEF1; however, more experimentation will be required to test this hypothesis. In addition, it is possible that this LEF1-Smad-R361complex does not act alone, and could include other transcription factors, such as β-catenin, to assist in downstream activation of WNT signaling. However, β-catenin binding has been understudied and independent publications yielded mixed results[163–165].

The four CMS subtypes differ in genetic and epigenetics, as well as the signaling pathways they follow. While CMS1 is characteristic of tumors with MSI, tumors with CIN can be subclassified on the basis of gene expression: CMS2 (canonical subtype), CMS3 (metabolic subtype) and CMS4 (mesenchymal subtype)[31]. Copy number variation in both oncogenes and tumor suppressor genes are found more frequently in CMS2 than in other subtypes and they display a strong upregulation of WNT and MYC downstream targets. Interestingly, SMAD4 mutations were profiled across different molecular subtypes and found to be most common with the CMS3 subtype[143], however, that classification was not specific to SMAD4 missense mutations. Based on our data and overall working model, we would expect that SMAD4 missense mutation (especially hotspot mutation R361) to be commonly found in the CMS2 subtype due to their marked upregulation of WNT pathway. However, this hypothesis could not be tested directly using our dataset due to the lack of RNA-Seq data in the MSKCC cohort. Nevertheless, based on our survival data, the interaction between Smad4-R361 and LEF/TCF protein complexes may be a good therapeutic target.

Our work raises additional questions that will be addressed in future studies. TGF-β not only induces Smad-mediated responses, but also activate Smad-independent responses, which allow additional versatility and diversification of TGF-β family responses. The non-Smad pathways include, but are not limited to, the activation of other mediators such as those involved in MAPK kinase pathways[39]. Loss of Smad4 has been shown to

induce tumorigenicity and 5-FU (5-fluoracil) resistance through activation of the Akt pathway, which results in upregulation of anti-apoptotic proteins[166]. Another study revealed that patients with advanced disease observed a markedly shorter progression-free survival time in patients with *SMAD4*-mutated tumors than in those wild-type for *SMAD4*, when treated with anti- epidermal growth factor receptor (EGFR) treatment[143]. This suggests that mutant Smad4 might be playing a role that affects signaling pathway, possibly via binding to other, yet to be identified transcription factors.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

OVERVIEW

As we summarize our findings and conclude our discussion, we end with an examination on future directions of this thesis project. We provide proposals for future experiments that will validate our findings, make suggestions of various avenues to pursue that could provide valuable insight, and outline how breakthroughs on defining this pathway further can open up possibilities to create better therapies for CRC cancer patients.

LIMITATIONS OF COMPLETED WORK & FUTURE DIRECTIONS

Elucidating molecular mechanism by which Smad4 R361 regulates WNT signaling

Here we proposed that Smad4-R361-mutant loses its tumor suppressive arm and retains binding to LEF1, resulting in enhanced WNT signaling. Although we have described this event in detail in this thesis, we still do not know exactly how this is occurring.

Firstly, we can speculate that the inability of Smad4mut to bind to Smad2/3 following TGF-β treatment increases the pool of available mutSmad4 to bind to LEF1. This would redirect Smad4mut to upregulate WNT signaling, while Smad4 wild-type regulates TGF-β and WNT signaling simultaneously. A possible way that we could test this hypothesis is to determine the effect of TGF-β signaling on WNT signaling in CRC cell lines. To test this, we could perform the TOP/FOP assay on the HCT116 3:6 cells and treat the cells with either the TGF-β soluble ligand or a TGFβ receptor kinase (ALK5/TGFβRI) inhibitor SB431542[167]. If our redirection hypothesis is true, we would expect that blocking TGF-β signaling would increase WNT signaling in our CRC cells

expressing both mutant and wild-type Smad4. However, there is a possibility that inhibition of TGF-β signaling will reduce WNT signaling, suggesting that the two pathways are working together. If this is our result, a way to possibly explain this would be through inhibitor Smad, Smad7, which is a downstream target gene of TGF-β signaling. In the cytoplasm, Smad7 is one of the points that link the two pathways. It is well known that Smad7 functions to compete with Smad2/3 for receptors, which leads to ubiquitination and degradation of receptors and hence inhibiting Smad2/3 phosphorylation and activation[38]. However, Smad7 has been found to disassemble the β-catenin destruction complex by binding to AXIN, thus stabilizing β-catenin and promoting its nuclear translocation[168]. When we treat parental HCT116 3:6 cells with TGF-β ligand, we have seen an upregulation of *SMAD7* transcript (data not shown). Given that Smad7 is upregulated in the presence of TGF-β signaling, and Smad7 has been shown to regulate WNT signaling; it would make sense that inhibiting TGF-β could also result in downregulation of WNT signaling.

In addition, a recent study by Voorneveld et al claims that BMP can affect WNT signaling depending on both SMAD4 and p53 status. In the normal colonic epithelium, the BMP pathway counteracts WNT signaling pathway in order to maintain tissue homeostasis[169]. In their study, they found that wild-type p53 is necessary for BMP to inhibit WNT signaling and loss of Smad4 completely reverses the BMP-WNT interaction, switching from inhibition to activation[164]. This is the opposite of what we see in CRC cells, as presence of Smad4 wild-type protein actually enhances WNT signaling compared to empty vector cells (expressing wild-type p53). Nevertheless, it would be interesting to explore a BMP-SMAD4-WNT axis to help us better understand the effects of Smad4 mutations on WNT signaling. If somehow BMP signaling is redirecting Smad4 to enhance WNT signaling, we could test this theory by using a similar experimental setup as proposed above. We would use our CRC cell lines and either treat with BMP ligand or

inhibit BMP signaling using an inhibitor (LDN-193189)[170], followed by a TOP/FOP luciferase assay to measure downstream WNT signaling. In general, understanding the molecular mechanisms of how Smad4mut is able to upregulate WNT signaling compared to wild-type Smad4 in colorectal cancer cells would further validate our findings stated in this body of work.

Uncovering Smad4 mutant phenotype *in vitro* and *in vivo*

Here, we presented evidence that mut-Smad4 is unable to promote canonical downstream TGF-β signaling (Figures 7 & 8). Although the work done in this thesis sheds light on the more novel role of Smad4 in colorectal cancer, understanding the phenotype of Smad4 mutations both *in vivo* and *in vivo* would provide invaluable information on Smad4mut functions in tissues.

TGF-β controls cell growth by activating cyclin dependent kinase inhibitors (CDKIs) such as p21 and p15 resulting in growth inhibition. Previous studies have demonstrated that TGF-β growth inhibition is dependent on the presence of Smad4[171]. In addition, restoration of Smad4-WT in both human and murine CRC cell lines restores TGF-β induced growth inhibition and reduces migration and invasion[172]. This, along the data presented in this thesis, would lead us to hypothesize that colorectal cancer cells expressing Smad4 mutants will exhibit higher tumorigenicity, migration, and invasion compared to their wild-type counterpart.

To determine the role of Smad4 mutations in CRC malignancy, we would use our SW480, SW620 and HCT116 3:6 cells and perform WST-1 and BrdU incorporation proliferation assays in the absence or presence of soluble TGF-β ligand. As a control, we could treat cells with TGF-β receptor kinase (ALK5/TGFβRI) inhibitor SB431542[167], which would abrogate TGF-β signaling by blocking the receptor. We have some preliminary data looking at cell migration and invasion in HCT116 3:6 cells (Fig 13A & B). HCT116 3:6 cells

expressing Smad4-R361 mutation seem to have an intermediate migratory and invasive phenotype compared to Smad4-null and Smad4-WT cells. In order to complement this experiment, we could also perform a cell migration assay (*in vitro* wound healing assay).

As mentioned in the introduction to this thesis, TGF-β can induce epithelial to mesenchymal transition (EMT) by inducing expression of EMT-promoting factors like Snail, Slug, ZEB and Twist[70]. As EMT has been shown to be dependent on the presence of Smad4[173], we could look into the effect of Smad4-mutant on inducing these EMT transcription factors. Investigating the details of how Smad4-mut expression is driving epithelial-mesenchymal transition can help us better understand the role in EMT and identifying how to prevent Smad4-mut from promoting this event in cancer patients.

Patients with mutant Smad4 have worse overall survival than patients with wild-type Smad4[143]. Our lab, as well as others, have extensively shown that downregulation of Smad4-WT through shRNA leads to a higher incidence of primary tumors and metastasis to the liver in mice[81,83]. Because all of the experiments proposed have been *in vitro*, we could look into how these Smad4-mut phenotypes translate in an *in vivo* mouse model. As orthotopic models accurately represent the progression and metastasis of CRC, we could use an orthotopic cecal injection to evaluate the effect of Smad4 mutations on CRC tumorigenicity and liver metastasis. Incidence of primary tumor formation would be assessed by looking at the colon, while incidence of metastasis would be assessed by looking in the liver. Taken together, these studies would allow us to understand whether Smad4-mut promote CRC progression and whether it would make sense to block Smad4-mut function in colorectal cancer patients.

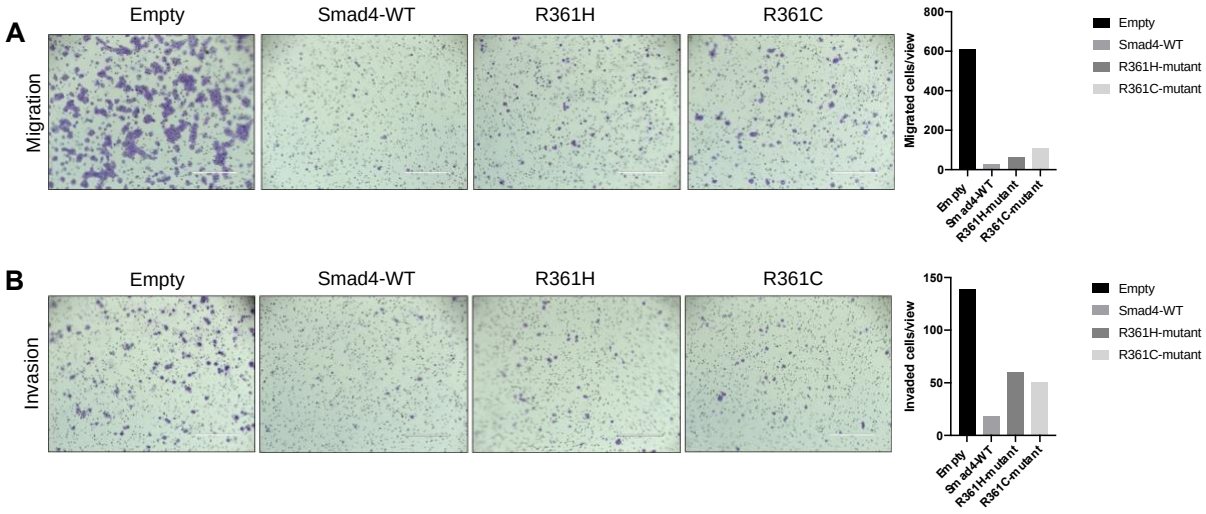

Figure 13. R361-mutant attenuates the inhibitor effect of Smad4 on cell migration and invasion in HCT116 3:6 cells. The imagines of HCT116 3:6 cells are representative of one pilot experiment (magnification 40X). Migration **(A)** and invasion **(B)** were determined using transwell assays. The result of the quantification of cell migration and invasion are shown (right).

Effects of Smad4 mutations on other pathways in the TGF-β superfamily

As previously mentioned, TGF-β can activate both Smad and non-Smad signaling pathways, such as Erk, JNK, and p38 MAPK kinase pathways[39]. Mehrvarz Sarshekeh et al's study revealed that patients with advanced disease observed a markedly shorter progression-free survival time in patients with *SMAD4*-mutated tumors than in those wild-type for *SMAD4* when treated with anti-epidermal growth factor receptor (EGFR) treatment[143]. The fact that EGFR like cetuximab and panitumumab[27,28] is provided to patients with wild-type *KRAS* (as any alterations in can lead to constitutive activation of EGFR and ultimately to drug resistance[28]) suggests that mut-Smad4 may be affecting other signaling pathways besides TGF-β.

To test this, we would need to use cell lines with wild type *KRAS* so that constitutive active signaling of the pathway does not interfere with the interpretation of our results. The tools to explore this, however, are limited given that most CRC cell lines harbor mutations in *KRAS* gene (a reflection of patient data, where over half of CRC patients harbor mutations in *KRAS*). There is one cell lines we could work with (COLO-320) that has wild-type KRAS and BRAF and baseline PI3K signaling. However, these cell lines express endogenous Smad4 protein. We would need to 1) generate COLO-320 isogenic cell lines using CRISPR-Cas9 to generate Smad4 KO clones and 2) re-introduce Smad4 retrovirally (as we did in HCT116 3:6 cells). Due to the limitation of cell lines, we could also use an organoid model system. By introducing inactivating mutations in APC, p53 and SMAD4 and activating KRAS(G12D) mutation, scientists have been able to use intestinal stem cells to better model colorectal cancer *in vitro* and *in vivo* [174,175]. In our case, we could use this system to generate p53, APC and SMAD4 null organoids via CRISPR-Cas9, then later introduce our Smad4 via retrovirus. As loss of SMAD4 has been shown to activate downstream Akt and Erk signaling[141,166,176], we can measure the levels of p-AKT and p-ERK in both Smad4-mut expressing cell lines and organoids. Overall, these observations

would allow us to understand the effect of Smad4 R361-mutations on both other signaling pathways in the TGF-β superfamily.

Investigating the mutSmad4-LEF1 axis in other cancers & diseases

It is worth investigating the mutSmad4-LEF1 signaling axis in other cancers, to confirm the mechanism we have uncovered is conserved among other tumor types. This will help us better understand how WNT signaling pathway could be used as a therapeutic option for patients who have abnormalities in this pathway.

Pancreatic ductal adenocarcinoma (PDAC) is one of the most aggressive cancers with a five-year survival of less than 5%[87]. Despite advancements in treatment and therapies, mortality rates for PDAC patients still remains high. Over half of PDAC patients with invasive disease have alterations in Smad4 protein[177], including deep deletions and missense mutations. Studies have shown that loss or inactivation of SMAD4 is associated with poor prognosis, due to an increased incidence of metastatic disease[178]. In addition, SMAD4 inactivating mutations occur in approximately 20% of all pancreatic cancer[92]. Similar to CRC tumors, they can occur in either MH1 or the MH2 domain, with the latter resulting in either loss of protein stability or disruption of dimerization ability of the Smads. One study has shown that missense mutations in SMAD4 at Y353C residue attenuates the inhibitory effect of SMAD4 on cell migration and invasion in PDAC cells without affecting cell proliferation[179]. This suggests that missense mutations in Smad4 could affect PDAC progression, just as we expect them to affect CRC progression.

According to cbioportal data, mutations in the R361 residue also are the most frequent type of Smad4 missense mutation found in PDAC patients, as evidenced by their lollipop plots (Figure 14A). In conclusion, understanding what the effect of R361-mutations on PDAC patients are, and the mechanism whereby they can affect PDAC progression can provide a firmer empirical basis for PDAC treatment.

Apart from cancers, it would also be interesting to understand the role of mutant Smad4 (and Smad4mut-LEF1 axis) in other diseases such as juvenile polyposis syndrome (JPS). Juvenile polyposis syndrome (JPS) is an autosomal disease that is characterized by the presence of benign polyps in the gastrointestinal tract, most commonly in the colon. Patients with JPS have up to a 50% risk of developing specific types of cancer, including colorectal (CRC) and other gastrointestinal (GI) malignancies[180,181]. Similarly, familial adenomatous polyposis (FAP) is a familial cancer predisposition syndrome with risk of multiple cancer types, including 100% risk of colon cancer absent intervention. The genetic basis of FAP is well-understood to be related to loss-of-function mutations, deletions, or duplications in the *APC* gene, known to hyperactivate the WNT pathway[182] This in turn leads to overexpression of the MYC oncogene and enhanced cell proliferation[183].

However, WNT pathway mutations have not been described in JPS. Instead, 30-40% of JPS cases are associated with germline frameshift and missense mutations in *SMAD4* and *BMPRIA* receptor[184]. In CRC, *SMAD4* pathway mutations (found with 10-35% frequency) are known to be associated with such “late” events. However, how *SMAD4* mutations can contribute to polyp formation and increased cancer risk in JPS patients remains unknown. Based on the work that we have done in our CRC cell lines, we would hypothesize that germline *SMAD4* missense mutations in JPS could promote polyp formation by increasing WNT signaling. To test this, we would use CRISPR-Cas9 technology to knockout endogenous Smad4 in normal colon organoids and immortalized colonocyte cell lines (NCM460, NCM365). We would then reconstitute them with naturally occurring SMAD4 JPS variants (focusing on Smad4-R361H/C mutations) using retroviral/lentiviral transduction. Once the JPS-associated Smad4 mutants (e.g., R361H/CS) are stably expressed, we would measure changes in WNT signaling by using WNT-reporter activity assay (TOP/FOP luciferase). In addition, we could also test for

different downstream signaling events correlated with WNT, such as activation of WNT target genes like Axin2 and β-catenin and enhanced cell proliferation. Finally, we can perform RNA-Seq on SMAD4-WT and SMAD4-mut cell lines and organoid, to compare and contrast their transcriptomes. Because we would that expect that the presence of *SMAD4* missense mutations would correlate with enhanced reliance on this signaling pathway, it would also be interesting to test the responses of patient derived organoids and JPS-model cell lines with existing WNT inhibitors such as PRI-724 and ICG-001. Our hypothesis is that mutations in Smad4 results in hyperactivation of WNT signaling, which could provide direct molecular targets for therapeutic intervention. Identification of such targets in JPS would be highly significant – and potentially translatable- development.

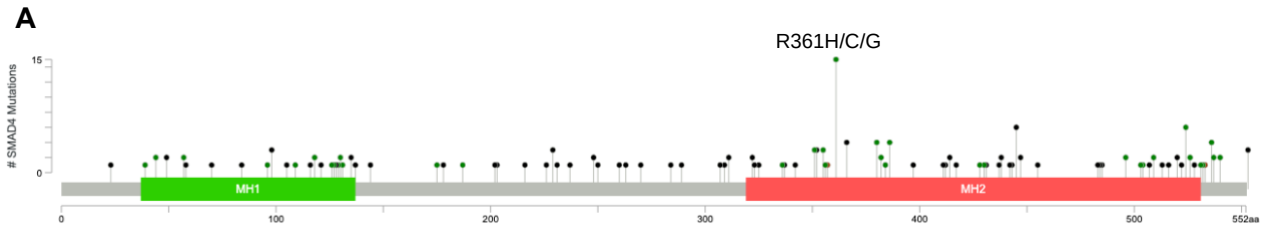

Figure 14. R361 mutation is prevalent in Smad4 in PDAC. (A) The prevalence and spectrum of SMAD4 mutations in multiple pancreatic adenocarcinoma studies from cbioportal.

Identification of other proteins that might interact with mutSmad4

Protein-protein interactions are essential to the function of all proteins and define their biological roles. Therefore, identification of protein binding partners can bring to light important information about protein functions. One of the most practical and high-yielding approaches is immunoprecipitation of a bait protein followed by mass spectrometry to identify co-precipitating proteins. This technique has been done to identify novel binding partners for mutations in other tumor suppressor genes, for example TP53. Some studies have used mass spectrometry to demonstrate that mutant p53, but not p53 wild type, are able to bind to novel proteins. For example, p53 R175H is able to bind to proteins Tim50 and MCM7[185] and p53 R273H can bind NRD1 specifically[186] . Thus, this method could be useful in identifying novel protein-protein interactions or targets of mutant Smad4 as well.

Although proteomics analysis has been done in CRC cells with either Smad4-null and Smad4 wild-type expressing cells[187], no such experiments have been done with cells expressing mutSmad4. To begin to investigate this, we could analyze the difference in protein binding between wild-type Smad4 and mutant Smad4 via mass spectrometry analysis. Because alteration at a single residue affects the binding of Smad4 to R-Smads, it would be interesting to know it this mutation allows for other (yet unidentified) proteins to bind to mutSmad4. If we wanted to further continue to study the role of mutSmad4 on WNT signaling, we could do an alternate version of the experiment. We could pretreat the cells with a WNT inhibitor and then send samples to mass spectrometry to identify proteins that might bind differently to mutSmad4 in a WNT dependent manner. Once candidates have been identified, we could test how the interaction of mutSmad4, and other proteins affect either WNT signaling or any other signaling pathway that we choose to continue to study. We could also test how the binding of these candidate proteins to mutSmad4 affects their phenotype both *in vitro* and *in vivo*, which could help us further define on why mutSmad4 (specifically R361 mutation) is so highly selected for in CRC cancer patients.

SUMMARY AND CONCLUDING REMARKS

The work presented here demonstrated a novel, previously unidentified, function for mutant Smad4 in CRC cells. We show that hotspot missense mutation in Smad4, R361H/C is unable to bind to R-Smads, Smad2/3 or accumulate in the nucleus following TGF-β treatment. In addition, Smad4-R361 mutants lose the ability to activate canonical downstream TGF-β signaling. Although initially these R361 mutations seemed to behave like loss of function mutations, we wanted to continue studying them given that they are so highly selected for in CRC tumors, which is more consistent with a gain-of- or neomorphic function. Following RNA-Sequencing, we were surprised to see that R361H has a gene signature associated with upregulation of the TCF/LEF protein complex and boosts downstream WNT signaling.

Having established that there is a gain of function role for Smad4 mutants, this information can be used to develop peptide-based inhibitors targeting the interacting domains between mutSmad4-LEF1, opening a therapeutic window for CRC patients harboring Smad4-R361 mutations.

REFERENCES

1. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin*. 2020;0(0):1-20. doi:10.3322/caac.21601
2. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology*. 2007;50(1):113-130. doi:10.1111/j.1365-2559.2006.02549.x
3. Division ERF. Molecular genetics of colorectal cancer. *Ann Gastroenterol*. 2011;27(1):9-14. doi:10.1146/annurev-pathol-011110-130235
4. Harris CC, Hollstein M. Prevention of Colorectal Cancer by Colonoscopic Polypectomy. *N Engl J Med*. 1993;329(18):1318-1327.
5. Church JM. Molecular Genetics of Colorectal Cancer. *Curr Ther Colon Rectal Surg*. 2005:343-347. doi:10.1016/B978-1-55664-480-1.50064-2
6. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell Press*. 1990;19(2):170-172.
7. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature*. 1998;396(6712):643-649. doi:10.1038/25292
8. Lin EI, Tseng LH, Gocke CD, et al. Mutational profiling of colorectal cancers with microsatellite instability. *Oncotarget*. 2015;6(39):42334-42344. doi:10.18632/oncotarget.5997
9. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet*. 2006;38(7):787-793. doi:10.1038/ng1834
10. De'angelis GL, Bottarelli L, Azzoni C, et al. Microsatellite instability in colorectal cancer. *Gastroenterology*. 2009;89(6):97-101. doi:10.23750/abm.v89i9-S.7960

11. Parsons R, Myeroff LL, Liu B, et al. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. *Cancer Res*. 1995;55(23):5548-5550. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7585632.

12. Singh MP, Rai S, Pandey A, Singh NK, Srivastava S. Molecular subtypes of colorectal cancer: An emerging therapeutic opportunity for personalized medicine. *Genes Dis*. 2019;(xxxx). doi:10.1016/j.gendis.2019.10.013

13. Pino MS, Chung DC. The Chromosomal Instability Pathway in Colon Cancer. *Gastroenterology*. 2010;138(6):2059-2072. doi:10.1053/j.gastro.2009.12.065

14. Schneikert J, Behrens J. The canonical Wnt signalling pathway and its APC partner in colon cancer development. *Gut*. 2007;56(3):417-425. doi:10.1136/gut.2006.093310

15. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. β-catenin is a target for the ubiquitin–proteasome pathway. *EMBO J*. 2011;8(2):148-164. doi:10.20886/jsek.2011.8.2.148-164

16. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. *Oncogene*. 2017;36(11):1461-1473. doi:10.1038/onc.2016.304

17. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal tumor development. *N Engl J Med*. 1988;319(9):525-532.

18. Eppert K, Scherer SW, Ozcelik H, et al. MADR2 maps to 18q21 and encodes a TGFβ-regulated MAD-related protein that is functionally mutated in colorectal carcinoma. *Cell*. 1996;86(4):543-552. doi:10.1016/S0092-8674(00)80128-2

19. Fearon ER, Cho KR, Nigro JM, et al. Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science (80-)*. 1990;247(4938):49-56. doi:10.1126/science.2294591

20. Samanta D, Datta PK. Alterations in the Smad pathway in human cancers. *Front Biosci*. 2012;17(4):1281-1293. doi:10.2741/3986

21. Yaeger R, Chatila WK, Lipsyc MD, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell*. 2018;33(1):125-136.e3. doi:10.1016/j.ccell.2017.12.004

22. Kang H, Jha S, Ivovic A, et al. Somatic SMAD3 -activating mutations cause melorheostosis by up-regulating the TGF- β / SMAD pathway. 2020;217(5).

23. Giehl K. Oncogenic Ras in tumour progression and metastasis. *Biol Chem*. 2005;386(3):193-205. doi:10.1515/BC.2005.025

24. Dhillon AS, Kolch W. Oncogenic B-Raf mutations: Crystal clear at last. *Cancer Cell*. 2004;5(4):303-304. doi:10.1016/S1535-6108(04)00087-X

25. Wymann MP, Björklöf K, Calvez R, et al. Phosphoinositide 3-kinase γ: A key modulator in inflammation and allergy. *Biochem Soc Trans*. 2003;31(1):275-280. doi:10.1042/bst0310275

26. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet*. 2019;394(10207):1467-1480. doi:10.1016/S0140-6736(19)32319-0

27. Porru M, Pompili L, Caruso C, Biroccio A, Leonetti C. Targeting kras in metastatic colorectal cancer: Current strategies and emerging opportunities. *J Exp Clin Cancer Res*. 2018;37(1):1-10. doi:10.1186/s13046-018-0719-1

28. Zhao B, Wang L, Qiu H, et al. Mechanisms of resistance to anti-EGFR therapy in colorectal cancer. *Oncotarget*. 2017;8(3):3980-4000. doi:10.18632/oncotarget.14012

29. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015;21(11):1350-1356. doi:10.1038/nm.3967

30. Loree JM, Pereira AAL, Lam M, et al. Classifying colorectal cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and

consensus molecular subtypes. *Clin Cancer Res*. 2018;24(5):1062-1072. doi:10.1158/1078-0432.CCR-17-2484

31. Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer*. 2017;17(2):79-92. doi:10.1038/nrc.2016.126
32. Menter DG, Davis JS, Broom BM, Overman MJ, Morris J, Kopetz S. Back to the Colorectal Cancer Consensus Molecular Subtype Future. *Curr Gastroenterol Rep*. 2019;21(2). doi:10.1007/s11894-019-0674-9
33. Zhang M, Wang M, Tan X, Li TF, Zhang YE, Chen D. Smad3 prevents β-catenin degradation and facilitates β-catenin nuclear translocation in chondrocytes. *J Biol Chem*. 2010;285(12):8703-8710. doi:10.1074/jbc.M109.093526
34. Ali Nawshad, Damian Medici2, Chang-Chih Liu and EDH. TGFβ3 inhibits E-cadherin gene expression in palate medial-edge epithelial cells through a Smad2-Smad4-LEF1 transcription complex. 2007;120(Pt 9):1646-1653. doi:10.1242/jcs.003129.TGF
35. Massagué J, Seoane J, Wotton D. Smad transcription factors. *Genes Dev*. 2005;19(23):2783-2810. doi:10.1101/gad.1350705
36. Attisano L, Lee-hoeflich ST. SMADs. *Encycl Ref Mol Pharmacol*. 2006:867-867. doi:10.1007/3-540-29832-0_1501
37. Massagué J. TGFb in cancer. *Cell*. 2008;134(2):215-230. doi:10.1016/j.cell.2008.07.001.TGF
38. Shi Y, Massagué J. Mechanisms of TGF-β signaling from cell membrane to the nucleus. *Cell*. 2003;113(6):685-700. doi:10.1016/S0092-8674(03)00432-X
39. Derynck R, Zhang YE. Smad-dependent and Smad-independent pathways in TGFB family signaling. 2003;4. doi:10.1038/nature02006
40. Goto K, Kamiya Y, Imamura T, Miyazono K, Miyazawa K. Selective inhibitory effects

of Smad6 on bone morphogenetic protein type I receptors. *J Biol Chem*. 2007;282(28):20603-20611. doi:10.1074/jbc.M702100200

41. Zhang S, Fei T, Zhang L, et al. Smad7 Antagonizes Transforming Growth Factor β Signaling in the Nucleus by Interfering with Functional Smad-DNA Complex Formation. *Mol Cell Biol*. 2007;27(12):4488-4499. doi:10.1128/mcb.01636-06
42. Shi Y, Wang YF, Jayaraman L, Yang H, Massagué J, Pavletich NP. Crystal structure of a Smad MH1 domain bound to DNA: Insights on DNA binding in TGF-β signaling. *Cell*. 1998;94(5):585-594. doi:10.1016/S0092-8674(00)81600-1
43. Zawel L, Dai J Le, Buckhaults P, et al. Human Smad3 and Smad4 Are Sequence-Specific Transcription Activators. 1998;1:611-617.
44. Yagi K, Goto D, Hamamoto T, Takenoshita S, Kato M, Miyazono K. Alternatively spliced variant of Smad2 lacking exon 3. *J Biol Chem*. 1999;274(2):703-709. doi:10.1074/jbc.274.2.703
45. Demagny H, Araki T, de Robertis EM. The tumor suppressor Smad4/DPC4 is regulated by phosphorylations that integrate FGF, Wnt, and TGF-β signaling. *Cell Rep*. 2014;9(2):688-700. doi:10.1016/j.celrep.2014.09.020
46. Koganti P, Levy-Cohen G, Blank M. Smurfs in protein homeostasis, signaling, and cancer. *Front Oncol*. 2018;8(AUG):1-11. doi:10.3389/fonc.2018.00295
47. Zhu H, Kavsak P, Abdollah S. A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation sequence identity and are most closely related to Pub1 , a ubiquitin. *Mol Biol*. 1999;400(August):687-693.
48. Bonni S, Wang HR, Causing CG, et al. TGF-β induces assembly of a Smad2-Smurf2 ubiquitin ligase complex that targets SnoN for degradation. *Nat Cell Biol*. 2001;3(6):587-595. doi:10.1038/35078562
49. Pierreux CE, Nicolas FJ, Hill CS. Transforming Growth Factor beta -Independent Shuttling of Smad4 between the Cytoplasm and Nucleus. *Mol Cell Biol*.

2000;20(23):9041-9054. doi:10.1128/mcb.20.23.9041-9054.2000

50. Qin B, Lam SSW, Lin K. Crystal structure of a transcriptionally active smad4 fragment. *Structure*. 1999;7(12):1493-1503. doi:10.1016/S0969-2126(00)88340-9

51. Feng X-H, Derynck R. Specificity and Versatility in Tgf-B Signaling Through Smads. *Annu Rev Cell Dev Biol*. 2005;21(1):659-693. doi:10.1146/annurev.cellbio.21.022404.142018

52. Wu JW, Fairman R, Penry J, Shi Y. Formation of a Stable Heterodimer between Smad2 and Smad4. *J Biol Chem*. 2001;276(23):20688-20694. doi:10.1074/jbc.M100174200

53. Shi Y, Hata A, Lo RS, Massagué J, Pavletich NP. A structural basis for mutational inactivation of the tumour suppressor Smad4. *Nature*. 1997;388(6637):87-93. doi:10.1038/40431

54. Wu JW, Hu M, Chai J, et al. Crystal Structure of a Phosphorylated Smad2: Recognition of Phosphoserine by the MH2 Domain and Insights on Smad Function in TGF-b Signaling Jia-Wei. *Mol Cell*. 2001;8(6):1277-1289. doi:10.1016/S1097-2765(01)00421-X

55. Fleming NI, Jorissen RN, Mouradov D, et al. SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer Res*. 2013;73(2):725-735. doi:10.1158/0008-5472.CAN-12-2706

56. Hill CS. Transcriptional control by the SMADs. *Cold Spring Harb Perspect Biol*. 2016;8(10). doi:10.1101/cshperspect.a022079

57. De Caestecker MP, Yahata T, Wang D, et al. The Smad4 activation domain (SAD) is a proline-rich, p300-dependent transcriptional activation domain. *J Biol Chem*. 2000;275(3):2115-2122. doi:10.1074/jbc.275.3.2115

58. Wu JW, Krawitz AR, Chai J, et al. Structural mechanism of Smad4 recognition by the nuclear oncoprotein Ski: Insights on Ski-mediated repression of TGF-β

signaling. *Cell*. 2002;111(3):357-367. doi:10.1016/S0092-8674(02)01006-1

59. Massagué J. TGFβ signalling in context. *Nat Rev Mol Cell Biol*. 2012;13(10):616-630. doi:10.1038/nrm3434

60. Hunt KK, Fleming JB, Abramian A, Zhang L, Evans DB, Chiao PJ. Overexpression of the tumor suppressor gene Smad4/DPC4 induces p21(waf1) expression and growth inhibition in human carcinoma cells. *Cancer Res*. 1998;58(24):5656-5661.

61. Seoane J, Le H-V, Shen L, Anderson SA, Massagué J. Integration of Smad and Forkhead Pathways in the Control of Neuroepithelial and Glioblastoma Cell Proliferation. *Cell*. 2004;117:211-223. doi:10.1016/s0065-2423(10)52003-1

62. Pardali K, Kurisaki A, Morén A, Ten Dijke P, Kardassis D, Moustakas A. Role of Smad proteins and transcription factor Sp1 in p21Waf1/Cip1 regulation by transforming growth factor-β. *J Biol Chem*. 2000;275(38):29244-29256. doi:10.1074/jbc.M909467199

63. Yoo YD, Choi JY, Lee SJ, et al. TGF-β-induced cell-cycle arrest through the p21(WAF1/CIP1)-G1 cyclin/CDKS-p130 pathway in gastric-carcinoma cells. *Int J Cancer*. 1999;83(4):512-517. doi:10.1002/(SICI)1097-0215(19991112)83:4<512::AID-IJC13>3.0.CO;2-Z

64. Alexandrow MG, Kawabata M, Aakre M, Moses HL. Overexpression of the c-Myc oncoprotein blocks the growth-inhibitory response but is required for the mitogenic effects of transforming growth factor β1. *Proc Natl Acad Sci U S A*. 1995;92(8):3239-3243. doi:10.1073/pnas.92.8.3239

65. Frederick JP, Liberati NT, Waddell DS, Shi Y, Wang X-F. Transforming Growth Factor β-Mediated Transcriptional Repression of c-myc Is Dependent on Direct Binding of Smad3 to a Novel Repressive Smad Binding Element. *Mol Cell Biol*. 2004;24(6):2546-2559. doi:10.1128/mcb.24.6.2546-2559.2004

66. Claassen GF, Hann SR. A role for transcriptional repression of p21CIP1 by c-Myc

in overcoming transforming growth factor β-induced cell-cycle arrest. *Proc Natl Acad Sci U S A*. 2000;97(17):9498-9503. doi:10.1073/pnas.150006697

67. Warner BJ, Blain SW, Seoane J, Massagué J. Myc Downregulation by Transforming Growth Factor β Required for Activation of the p15Ink4b G1 Arrest Pathway. *Mol Cell Biol*. 1999;19(9):5913-5922. doi:10.1128/mcb.19.9.5913

68. Yeh HW, Lee SS, Chang CY, Lang YD, Jou YS. A new switch for TGFβ in cancer. *Cancer Res*. 2019;79(15):3797-3805. doi:10.1158/0008-5472.CAN-18-2019

69. Calon A, Espinet E, Palomo-Ponce S, et al. Dependency of colorectal cancer on a TGF-beta-driven programme in stromal cells for metastasis initiation. *Cancer Cell*. 2012;22(5):571-584. doi:10.1016/j.ccr.2012.08.013

70. Koen EJ, Collier AB. The basics of epithelial-mesenchymal transition. *Phys Plasmas*. 2010;To be subm(May):1420-1428. doi:10.1172/JCI39104.1420

71. Moustakas A, Heldin CH. Induction of epithelial-mesenchymal transition by transforming growth factor β. *Semin Cancer Biol*. 2012;22(5-6):446-454. doi:10.1016/j.semcancer.2012.04.002

72. Runyan CE, Hayashida T, Hubchak S, Curley JF, Schnaper HW. Role of SARA (SMAD anchor for receptor activation) in maintenance of epithelial cell phenotype. *J Biol Chem*. 2009;284(37):25181-25189. doi:10.1074/jbc.M109.032847

73. Hesling C, Fattet L, Teyre G, et al. Antagonistic regulation of EMT by TIF1γ and Smad4 in mammary epithelial cells. *EMBO Rep*. 2011;12(7):665-672. doi:10.1038/embor.2011.78

74. Ozdamar B, Bose R, Barrios-Rodiles M, Wang HR, Zhang Y, Wrana JL. Regulation of the polarity protein Par6 by TGFβ receptors controls epithelial cell plasticity. *Science (80-)*. 2005;307(5715):1603-1609. doi:10.1126/science.1105718

75. Viloria-Petit AM, Wrana JL. The TGFβ-Par6 polarity pathway: Linking the Par complex to EMT and breast cancer progression. *Cell Cycle*. 2010;9(4):623-624.

doi:10.4161/cc.9.4.10779

76. Lamouille S, Derynck R. Cell size and invasion in TGF-β-induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *J Cell Biol*. 2007;178(3):437-451. doi:10.1083/jcb.200611146

77. Lamouille S, Connolly E, Smyth JW, Akhurst RJ, Derynck R. TGf-β-induced activation of mTOR complex 2 drives epithelial-mesenchymal transition and cell invasion. *Development*. 2012;139(10):1259-1273. doi:10.1242/jcs.095299

78. Jain RK. Normalization of tumor vasculature: An emerging concept in antiangiogenic therapy. *Science (80-)*. 2005;307(5706):58-62. doi:10.1126/science.1104819

79. Folkman J. Tumor Angiogenesis: Therapeutic Implications. *N Engl J Med*. 1971;262:1182-1186. doi:10.1056/NEJM196006162622401

80. Zetter BR. Angiogenesis and tumor metastasis. *Annu Rev Med*. 1998;49(1):407-424. doi:10.1146/annurev.med.49.1.407

81. Dews M, Tan GS, Hultine S, et al. Masking epistasis between MYC and TGF-β pathways in antiangiogenesis-mediated colon cancer suppression. *J Natl Cancer Inst*. 2014;106(4). doi:10.1093/jnci/dju043

82. Papageorgis P, Cheng K, Ozturk S, et al. Smad4 inactivation promotes malignancy and drug resistance of colon cancer. *Cancer Res*. 2011;71(3):998-1008. doi:10.1158/0008-5472.CAN-09-3269

83. Alazzouzi H, Alhopuro P, Salovaara R, et al. SMAD4 as a prognostic marker in colorectal cancer. *Clin Cancer Res*. 2005;11(7):2606-2611. doi:10.1158/1078-0432.CCR-04-1458

84. Losi L, Bouzourene H, Benhattar J. Loss of Smad4 expression predicts liver metastasis in human colorectal cancer. *Oncol Rep*. 2007;17(5):1095-1099.

85. Jia X, Shanmugam C, Paluri RK, et al. Prognostic value of loss of heterozygosity

and sub-cellular localization of SMAD4 varies with tumor stage in colorectal cancer. *Oncotarget*. 2017;8(12):20198-20212. doi:10.18632/oncotarget.15560

86. Woodford-Richens KL, Rowan AJ, Poulsom R, et al. Comprehensive analysis of SMAD4 mutations and protein expression in juvenile polyposis: Evidence for a distinct genetic pathay and polyp morphology in SMAD4 mutation caries. *Am J Pathol*. 2001;159(4):1293-1300. doi:10.1016/S0002-9440(10)62516-3

87. Ahmed S, Bradshaw A-D, Gera S, Dewan M, Xu R. The TGF-β/Smad4 Signaling Pathway in Pancreatic Carcinogenesis and Its Clinical Significance. *J Clin Med*. 2017;6(1):5. doi:10.3390/jcm6010005

88. Rowan AJ, Gorman P, Halford S, et al. SMAD4 mutations in colorectal cancer probably occur before chromosomal instability , but after divergence of the microsatellite instability pathway. 2001;98(17):9719-9723.

89. Maurice D, Pierreux CE, Howell M, Wilentz RE, Owen MJ, Hill CS. Loss of Smad4 function in pancreatic tumors: C-terminal truncation leads to decreased stability. *J Biol Chem*. 2001;276(46):43175-43181. doi:10.1074/jbc.M105895200

90. De Bosscher K, Hill CS, Nicolás FJ. Molecular and functional consequences of Smad4 C-terminal missense mutations in colorectal tumour cells. *Biochem J*. 2004;379(1):209-216. doi:10.1042/BJ20031886

91. Wan M, Huang J, Jhala NC, et al. SCFβ-TrCP1 controls Smad4 protein stability in pancreatic cancer cells. *Am J Pathol*. 2005;166(5):1379-1392. doi:10.1016/S0002-9440(10)62356-5

92. Xu J, Attisano L. Mutations in the tumor suppressors Smad2 and Smad4 inactivate transforming growth factor β signaling by targeting Smads to the ubiquitin-proteasome pathway. *Proc Natl Acad Sci U S A*. 2000;97(9):4820-4825. doi:10.1073/pnas.97.9.4820

93. Miyaki M, Iijima T, Konishi M, et al. Higher frequency of Smad4 gene mutation in

human colorectal cancer with distant metastasis. *Oncogene*. 1999;18(20):3098-3103. doi:10.1038/sj.onc.1202642

94. Miyaki M, Kuroki T. Role of Smad4 (DPC4) inactivation in human cancer. *Biochem Biophys Res Commun*. 2003;306(4):799-804. doi:10.1016/S0006-291X(03)01066-0

95. Thiagalingam S, Lengauer C, Leach FS, et al. Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. *Nat Genet*. 1996;13(3):343-346. doi:10.1038/ng0796-343

96. Clevers H. Wnt/β-Catenin Signaling in Development and Disease. *Cell*. 2006;127(3):469-480. doi:10.1016/j.cell.2006.10.018

97. Verheyen EM, Gottardi CJ. Regulation of Wnt/β-catenin signaling by protein kinases. *Dev Dyn*. 2010;239(1):34-44. doi:10.1002/dvdy.22019

98. Amit S, Amit S, Hatzubai A, et al. Axin-mediated CKI phosphorylation of β-catenin at Ser 45: a molecular switch for the Wnt pathway. *Genes Dev*. 2002;16:1066-1076. doi:10.1101/gad.230302.somal

99. Kitagawa M, Hatakeyama S, Shirane M, et al. An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of β-catenin. *EMBO J*. 1999;18(9):2401-2410. doi:10.1093/emboj/18.9.2401

100. Hart M, Concordet JP, Lassot I, et al. The F-box protein β-TrCP associates with phosphorylated β-catenin and regulates its activity in the cell. *Curr Biol*. 1999;9(4):207-211. doi:10.1016/S0960-9822(99)80091-8

101. Latres E, Chiaur DS, Pagano M. The human F box protein β-Trcp associates with the Cul1/Skp1 complex and regulates the stability of β-catenin. *Oncogene*. 1999;18(4):849-854. doi:10.1038/sj.onc.1202653

102. Chan EF, Gat U, McNiff JM, Fuchs E. A common human skin tumour is caused by activating mutations in β- catenin. *Nat Genet*. 1999;21(4):410-413.

doi:10.1038/7747

103. Kim S, Jeong S. Mutation hotspots in the β-catenin gene: Lessons from the human cancer genome databases. *Mol Cells*. 2019;42(1):8-16. doi:10.14348/molcells.2018.0436

104. Gao C, Wang Y, Broaddus R, Sun L, Xue F, Zhang W. Exon 3 mutations of CTNNB1 drive tumorigenesis: A review. *Oncotarget*. 2018;9(4):5492-5508. doi:10.18632/oncotarget.23695

105. Behrens J, Von Kries JP, Kühl M, et al. Functional interaction of β-catenin with the transcription factor LEF- 1. *Nature*. 1996;382(6592):638-642. doi:10.1038/382638a0

106. Arce L, Yokoyama NN, Waterman ML. Diversity of LEF/TCF action in development and disease. *Oncogene*. 2006;25(57):7492-7504. doi:10.1038/sj.onc.1210056

107. Kinzler KW, Vogelstein B. Lessons from Hereditary Colorectal Cancer. *Cell*. 1996;87:159-170. doi:10.1016/S0092-8674(00)81333-1

108. Kitaeva MN, Grogan L, Williams JP, et al. Mutations in β-catenin are uncommon in colorectal cancer occurring in occasional replication error-positive tumors. *Cancer Res*. 1997;57(20):4478-4481.

109. Polakis P. The oncogenic activation of β-catenin. *Curr Opin Genet Dev*. 1999;9(1):15-21. doi:10.1016/S0959-437X(99)80003-3

110. Cadigan KM, Waterman ML. TCF / LEFs and Wnt Signaling in the Nucleus. *Cold Spring Harb Perspect Biol*. 2012:1-22.

111. Arce L, Pate KT, Waterman ML. Groucho binds two conserved regions of LEF-1 for HDAC-dependent repression. *BMC Cancer*. 2009;9:1-14. doi:10.1186/1471-2407-9-159

112. Cavallo RA, Cox RT, Moline MM, et al. Drosophila Tcf and Groucho interact to repress wingless signalling activity. *Nature*. 1998;395(6702):604-608.

doi:10.1038/26982

113. Daniels DL, Weis WI. β-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol*. 2005;12(4):364-371. doi:10.1038/nsmb912

114. Hecht A. The p300/CBP acetyltransferases function as transcriptional coactivators of beta-catenin in vertebrates. *EMBO J*. 2000;19(8):1839-1850. doi:10.1093/emboj/19.8.1839

115. Teo JL, Kahn M. The Wnt signaling pathway in cellular proliferation and differentiation: A tale of two coactivators. *Adv Drug Deliv Rev*. 2010;62(12):1149-1155. doi:10.1016/j.addr.2010.09.012

116. Tsedensodnom O, Koga H, Rosenberg S, et al. Identification of T-Cell Factor-4 isoforms that contribute to the malignant phenotype of hepatocellular carcinoma cells. *Exp Cell Res*. 2012;23(1):1-7. doi:10.1038/jid.2014.371

117. Tomimaru Y, Koga H, Shin TH, Xu CQ, Wands JR, Kim M. The SxxSS motif of T-cell factor-4 isoforms modulates Wnt/β-catenin signal activation in hepatocellular carcinoma cells. *Cancer Lett*. 2013;336(2):359-369. doi:10.1016/j.canlet.2013.03.031

118. Santiago L, Daniels G, Wang D, Deng FM, Lee P. Wnt signaling pathway protein LEF1 in cancer, as a biomarker for prognosis and a target for treatment. *Am J Cancer Res*. 2017;7(6):1389-1406.

119. Waterman ML. Lymphoid enhancer factor T cell factor expression in colorectal cancer (Differential Expression of TCFs in tissues & organs).pdf. 2004:41-52.

120. Wang WJ, Yao Y, Jiang LL, et al. Increased LEF1 expression and decreased notch2 expression are strong predictors of poor outcomes in colorectal cancer patients. *Dis Markers*. 2013;35(5):395-405. doi:10.1155/2013/983981

121. Su MC, Chen CT, Huang FI, Chen YL, Jeng YM, Lin CY. Expression of LEF1 is an

independent prognostic factor for patients with oral squamous cell carcinoma. *J Formos Med Assoc*. 2014;113(12):934-939. doi:10.1016/j.jfma.2013.07.012

122. Bleckmann A, Siam L, Klemm F, et al. Nuclear LEF1/TCF4 correlate with poor prognosis but not with nuclear β-catenin in cerebral metastasis of lung adenocarcinomas. *Clin Exp Metastasis*. 2013;30(4):471-482. doi:10.1007/s10585-012-9552-7
123. Wang WJ, Yao Y, Jiang LL, et al. Knockdown of Lymphoid Enhancer factor 1 Inhibits Colon Cancer Progression In Vitro and In Vivo. *PLoS One*. 2013;8(10):1-11. doi:10.1371/journal.pone.0076596
124. Gao X, Mi Y, Ma Y, Jin W. LEF1 regulates glioblastoma cell proliferation, migration, invasion, and cancer stem-like cell self-renewal. *Tumor Biol*. 2014;35(11):11505-11511. doi:10.1007/s13277-014-2466-z
125. Colon Cancer Treatment By Stage. http://www.webmd.com/colorectal-cancer/guide/treatment-stage?page=2.
126. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. *Nat Rev Dis Prim*. 2015;1. doi:10.1038/nrdp.2015.65
127. Vatandoust S, Price TJ, Karapetis CS. Colorectal cancer: Metastases to a single organ. *World J Gastroenterol*. 2015;21(41):11767-11776. doi:10.3748/wjg.v21.i41.11767
128. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. *Nat Rev Dis Prim*. 2015;1:1-25. doi:10.1038/nrdp.2015.65
129. NIH. Targeted Cancer Therapies. https://www.cancer.gov/about-cancer/treatment/types/targeted-therapies/targeted-therapies-fact-sheet. Accessed September 11, 2020.
130. Xie Y, Chen Y, Fang J. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct Target Ther*. 2020;(December 2019).

doi:10.1038/s41392-020-0116-z

131. Huynh LK, Hipolito CJ, Dijke P Ten. A perspective on the development of TGF-β inhibitors for cancer treatment. *Biomolecules*. 2019;9(11). doi:10.3390/biom9110743

132. Tolcher AW, Berlin JD, Cosaert J, et al. A phase 1 study of anti-TGFβ receptor type-II monoclonal antibody LY3022859 in patients with advanced solid tumors. *Cancer Chemother Pharmacol*. 2017;79(4):673-680. doi:10.1007/s00280-017-3245-5

133. Morris JC, Tan AR, Olencki TE, et al. Phase I study of GC1008 (Fresolimumab): A human anti-transforming growth factor-beta (TGFβ) monoclonal antibody in patients with advanced malignant melanoma or renal cell carcinoma. *PLoS One*. 2014;9(3). doi:10.1371/journal.pone.0090353

134. Nicholas Dias, Yung Peng RK. Wnt Signaling and Colorectal Cancer. *Physiol Behav*. 2017;176(3):139-148. doi:10.1016/j.physbeh.2017.03.040

135. Gurney A, Axelrod F, Bond CJ, et al. Wnt pathway inhibition via the targeting of Frizzled receptors results in decreased growth and tumorigenicity of human tumors. *Proc Natl Acad Sci U S A*. 2012;109(29):11717-11722. doi:10.1073/pnas.1120068109

136. Le PN, McDermott JD, Jimeno A. Targeting the Wnt pathway in human cancers: Therapeutic targeting with a focus on OMP-54F28. *Pharmacol Ther*. 2015;146:1-11. doi:10.1016/j.pharmthera.2014.08.005

137. Fevr T, Robine S, Louvard D, Huelsken J. Wnt/β-Catenin Is Essential for Intestinal Homeostasis and Maintenance of Intestinal Stem Cells. *Mol Cell Biol*. 2007;27(21):7551-7559. doi:10.1128/mcb.01034-07

138. Emami KH, Nguyen C, Ma H, et al. A small molecule inhibitor of protein transcription β-catenin/CREB-binding. *Proc Natl Acad Sci U S A*. 2004;101(47):12682-12687. doi:10.1073/pnas.0404875101

139. Seoane J, Gomis RR. TGF-β family signaling in tumor suppression and cancer progression. *Cold Spring Harb Perspect Biol*. 2017;9(12). doi:10.1101/cshperspect.a022277

140. TCGA Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487(7407):330-337. doi:10.1038/nature11252

141. Papageorgis P, Cheng K, Ozturk S, et al. Smad4 inactivation promotes malignancy and drug resistance of colon cancer. *Cancer Res*. 2011;71(3):998-1008. doi:10.1158/0008-5472.CAN-09-3269

142. Iacobuzio-Donahue CA, Song J, Parmiagiani G, Yeo CJ, Hruban RH, Kern SE. Missense Mutations of MADH4: Characterization of the Mutational Hot Spot and Functional Consequences in Human Tumors. *Clin Cancer Res*. 2004;10(5):1597-1604. doi:10.1158/1078-0432.CCR-1121-3

143. Sarshekeh AM, Advani S, Overman MJ, et al. Association of SMAD4 mutation with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer. *PLoS One*. 2017;12(3). doi:10.1371/journal.pone.0173345

144. Kim E, Ilic N, Shrestha Y, et al. Systematic functional interrogation of rare cancer variants identifies oncogenic alleles. *Cancer Discov*. 2016;6(7):714-726. doi:10.1158/2159-8290.CD-16-0160

145. Schwarte-Waldhoff I, Volpert O V., Bouck NP, et al. Smad4/DPC4-mediated tumor suppression through suppression of angiogenesis. *Proc Natl Acad Sci U S A*. 2000;97(17):9624-9629. doi:10.1073/pnas.97.17.9624

146. Morén A, Hellman U, Inada Y, Imamura T, Heldin CH, Moustakas A. Differential ubiquitination defines the functional status of the tumor suppressor Smad4. *J Biol Chem*. 2003;278(35):33571-33582. doi:10.1074/jbc.M300159200

147. Brannon AR, Vakiani E, Sylvester BE, et al. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic

colorectal cancer lesions. *Genome Biol*. 2014;15(8). doi:10.1186/s13059-014-0454-7

148. Ahmed D, Eide PW, Eilertsen IA, et al. Epigenetic and genetic features of 24 colon cancer cell lines. 2013;(0424). doi:10.1038/oncsis.2013.35

149. Koi M, Umar A, Chauhan DP, et al. Human Chromosome 3 Corrects Mismatch Repair Deficiency and Microsatellite Instability and Reduces N-Methyl-N'-nitro-N-nitrosoguanidine Tolerance in Colon Tumor Cells with Homozygous hMLHl Mutation. *Cancer Res*. 1994;54(16):4308-4312.

150. Dennler S, Itoh S, Vivien D, Dijke P Ten, Huet S, Gauthier JM. Direct binding of Smad3 and Smad4 to critical TGFβ-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J*. 1998;17(11):3091-3100. doi:10.1093/emboj/17.11.3091

151. Miyazono K, Maeda S, Imamura T. BMP receptor signaling: Transcriptional targets, regulation of signals, and signaling cross-talk. *Cytokine Growth Factor Rev*. 2005;16(3 SPEC. ISS.):251-263. doi:10.1016/j.cytogfr.2005.01.009

152. Yokoyama Y, Watanabe T, Tamura Y, Hashizume Y, Miyazono K, Ehata S. Autocrine BMP-4 signaling is a therapeutic target in colorectal cancer. *Cancer Res*. 2017;77(15):4026-4038. doi:10.1158/0008-5472.CAN-17-0112

153. Bao Y, Guo Y, Yang Y, et al. PRSS8 suppresses colorectal carcinogenesis and metastasis. *Oncogene*. 2019;38(4):497-517. doi:10.1038/s41388-018-0453-3

154. Goldenring JR, Nam KT. Rab25 as a tumour suppressor in colon carcinogenesis. *Br J Cancer*. 2011;104(1):33-36. doi:10.1038/sj.bjc.6605983

155. Liu S, Fan W, Gao X, et al. Estrogen receptor alpha regulates the Wnt/β-catenin signaling pathway in colon cancer by targeting the NOD-like receptors. *Cell Signal*. 2019;61(April):86-92. doi:10.1016/j.cellsig.2019.05.009

156. Labbe E, Letamendia A, Attisano L. Association of Smads with lymphoid enhancer

binding factor 1/T cell-specific factor mediates cooperative signaling by the transforming growth factor-β and Wnt pathways. *Proc Natl Acad Sci U S A*. 2000;97(15):8358-8363. doi:10.1073/pnas.150152697

157. Qu H, Zheng L, Jiao W, et al. Smad4 suppresses the tumorigenesis and aggressiveness of neuroblastoma through repressing the expression of heparanase. *Sci Rep*. 2016;6(February):1-14. doi:10.1038/srep32628
158. Yoshioka Y, Togashi Y, Chikugo T, et al. Clinicopathological and genetic differences between low-grade and high-grade colorectal mucinous adenocarcinomas. *Cancer*. 2015;121(24):4359-4368. doi:10.1002/cncr.29676
159. Payne SR, Kemp CJ. Tumor suppressor genetics. *Carcinogenesis*. 2005;26(12):2031-2045. doi:10.1093/carcin/bgi223
160. Arai T, Akiyama Y, Yamamura A, et al. Allelotype analysis of early colorectal cancers with lymph node metastasis. *Int J Cancer*. 1998;79(4):418-423. doi:10.1002/(SICI)1097-0215(19980821)79:4<418::AID-IJC18>3.0.CO;2-0
161. Martinez-Lopez E, Abad A, Font A, et al. Allelic loss on chromosome 18q as a prognostic marker in stage II colorectal cancer. *Gastroenterology*. 1998;114(6):1180-1187. doi:10.1016/S0016-5085(98)70423-8
162. Lim SK, Hoffmann FM. Smad4 cooperates with lymphoid enhancer-binding factor 1/T cell-specific factor to increase c-myc expression in the absense of TGF-β signaling. *Proc Natl Acad Sci U S A*. 2006;103(49):18580-18585. doi:10.1073/pnas.0604773103
163. Tian X, Zhang J, Tan TK, et al. Association of β-catenin with P-Smad3 but not LEF-1 dissociates in vitro profibrotic from anti-inflammatory effects of TGF-β1. *J Cell Sci*. 2013;126(1):67-76. doi:10.1242/jcs.103036
164. Voorneveld PW, Kodach LL, Jacobs RJ, et al. The BMP pathway either enhances or inhibits the Wnt pathway depending on the SMAD4 and p53 status in CRC. *Br J*

Cancer. 2015;112(1):122-130. doi:10.1038/bjc.2014.560

165. Calvo-Sánchez MI, Fernández-Martos S, Carrasco E, et al. A role for the Tgf- β /Bmp co-receptor Endoglin in the molecular oscillator that regulates the hair follicle cycle. *J Mol Cell Biol*. 2018;11(1):39-52. doi:10.1093/jmcb/mjy051

166. Zhang B, Chen X, Bae S, Singh K, Washington MK, Datta PK. Loss of Smad4 in colorectal cancer induces resistance to 5-fluorouracil through activating Akt pathway. *Br J Cancer*. 2014;110(4):946-957. doi:10.1038/bjc.2013.789

167. Inman GJ, Nicolás FJ, Callahan JF, et al. SB-431542 is a potent and specific inhibitor of transforming growth factor-β superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol Pharmacol*. 2002;62(1):65-74. doi:10.1124/mol.62.1.65

168. Tang Y, Liu Z, Zhao L, Clemens TL, Cao X. Smad7 stabilizes β-catenin binding to E-cadherin complex and promotes cell-cell adhesion. *J Biol Chem*. 2008;283(35):23956-23963. doi:10.1074/jbc.M800351200

169. Radtke F, Clevers H. Self-renewal and cancer of the gut: Two sides of a coin. *Science (80-)*. 2005;307(5717):1904-1909. doi:10.1126/science.1104815

170. Sanvitale CE, Kerr G, Chaikuad A, et al. A New Class of Small Molecule Inhibitor of BMP Signaling. *PLoS One*. 2013;8(4). doi:10.1371/journal.pone.0062721

171. Pardali K, Kurisaki A, Morén A, Ten Dijke P, Kardassis D, Moustakas A. Role of Smad proteins and transcription factor Sp1 in p21Waf1/Cip1 regulation by transforming growth factor-β. *J Biol Chem*. 2000;275(38):29244-29256. doi:10.1074/jbc.M909467199

172. Zhang B, Halder SK, Kashikar ND, et al. Antimetastatic Role of Smad4 Signaling in Colorectal Cancer. *Gastroenterology*. 2010;138(3). doi:10.1053/j.gastro.2009.11.004

173. Zhao M, Mishra L, Deng CX. The role of TGF-β/SMAD4 signaling in cancer. *Int J*

Biol Sci. 2018;14(2):111-123. doi:10.7150/ijbs.23230

174. Matano M, Date S, Shimokawa M, et al. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat Med*. 2015;21(3):256-262. doi:10.1038/nm.3802

175. Drost J, Van Jaarsveld RH, Ponsioen B, et al. Sequential cancer mutations in cultured human intestinal stem cells. *Nature*. 2015;521(7550):43-47. doi:10.1038/nature14415

176. Zhang B, Leng C, Wu C, et al. Smad4 sensitizes colorectal cancer to 5-fluorouracil through cell cycle arrest by inhibiting the PI3K/Akt/CDC2/survivin cascade. *Oncol Rep*. 2016;35(3):1807-1815. doi:10.3892/or.2015.4479

177. Ali S, Choen C, Little J, Sequeira J, Mosunjac M, Siddiqui M. The Utility of SMAD4 as a Diagnostic Immunohistochemical Marker for Pancreatic Adenocarcinoma, and its Expression in Other Solid Tumors. *Diagn Cytopathol*. 2007;35(10):644-648. doi:10.1002/dc

178. Ahmed S, Schwartz C, Dewan M, Xu R. The Promising Role of TGF- β/SMAD4 in Pancreatic Cancer: The future targeted therapy. *J Cancer Treat Diagnosis*. 2019;3(2):1-7. doi:10.29245/2578-2967/2019/2.1141

179. Wang Z, Li Y, Zhan S, et al. SMAD4 Y353C promotes the progression of PDAC. *BMC Cancer*. 2019;19(1):1-12. doi:10.1186/s12885-019-6251-7

180. Cichy W, Klincewicz B, Plawski A. Juvenile polyposis syndrome. *Arch Med Sci*. 2014;10(3):570-577. doi:10.5114/aoms.2014.43750

181. Coburn MC, Pricolo VE, DeLuca FG, Bland KI. Malignant potential in intestinal juvenile polyposis syndromes. *Ann Surg Oncol*. 1995;2(5):386-391.

182. Nieuwenhuis MH, Vasen HFA. Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): A review of the literature. *Crit Rev Oncol Hematol*. 2007;61(2):153-161. doi:10.1016/j.critrevonc.2006.07.004

183. He TC, Sparks AB, Rago C, et al. Identification of c-MYC as a target of the APC pathway. *Science (80-)*. 1998;281(5382):1509-1512. doi:10.1126/science.281.5382.1509

184. Boettcher S, Miller PG, Sharma R, et al. A dominant-negative effect drives selection of TP53 missense mutations in myeloid malignancies. *Science (80-)*. 2019;365(6453):599-604. doi:10.1126/science.aax3649

185. Deb S, Graves PR. Identification of Novel Mutant p53 Interacting Proteins by Proteomic Analysis. 962:85-94. doi:10.1007/978-1-62703-236-0

186. Coffill CR, Muller PAJ, Oh HK, et al. Mutant p53 interactome identifies nardilysin as a p53R273H-specific binding partner that promotes invasion. *EMBO Rep*. 2012;13(7):638-644. doi:10.1038/embor.2012.74

187. Ali NA, McKay MJ, Molloy MP. Proteomics of Smad4 regulated transforming growth factor-beta signalling in colon cancer cells. *Mol Biosyst*. 2010;6(11):2332-2338. doi:10.1039/c0mb00016g